



The
Patent
Office

GB9 2241

09/743,647

#18



INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

27/01
459

27/01

5

REC'D 15 NOV 1999

WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

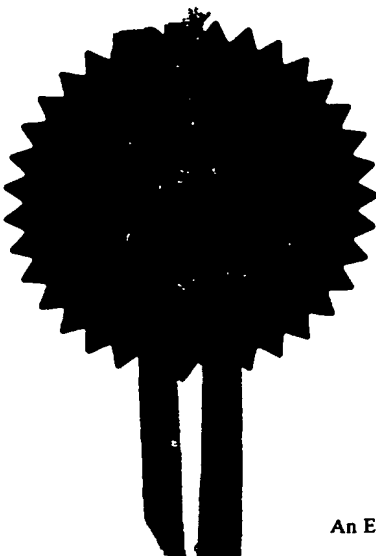
Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

BEST AVAILABLE COPY

Signed *Andrew Gersey*

Dated 5 November 1999

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)



An Executive Agency of the Department of Trade and Industry

G 912241
1/77**Request for grant of a patent***(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH15 JUL 98 03:57:27-10 002980
P01/7700 25.00 - 951535-1

1.	Your reference	SCB/50965/000		
2.	Patent application number <i>(The Patent Office will fill in this part)</i>	14 JUL 1998	9815284.6	
3.	Full name, address and postcode of the or of each applicant <i>(underline all surnames)</i>	JANSSEN PHARMACEUTICA N.V. Turnhoutseweg 30 B-2340 BEERSE BELGIUM		
	Patents ADP number <i>(if you know it)</i>	53193900 /		
	If the applicant is a corporate body, give the country/state of its incorporation	BELGIUM		
4.	Title of the invention	CLONING AND CHARACTERISATION OF NOVEL MAMMALIAN PEPTIDASES		
5.	Name of your agent <i>(if you have one)</i>	BOULT WADE TENNANT 27 FURNIVAL STREET LONDON EC4A 1PQ		
	"Address for service" in the United Kingdom to which all correspondence should be sent <i>(including the postcode)</i>	42001		
	Patents ADP number <i>(if you know it)</i>			
	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and <i>(if you know it)</i> the or each application number	Country	Priority application number <i>(if you know it)</i>	Date of filing <i>(day/month/year)</i>
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing <i>(day / month / year)</i>	
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? <i>(Answer 'Yes' if:</i> a) any applicant named in part 3 is not an inventor, or b) there is an inventor who is not named as an applicant, or c) any named applicant is a corporate body. See note (d))	YES		

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form -

Description 63

Claim(s) 6

Abstract -

Drawing(s) 16 + 16

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents
(Please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Date

Boulet Wade Tennant

14 July 1998

12. Name and daytime telephone number of person to contact in the United Kingdom **COLM D. MURPHY**
0171 404 5921

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) For details of the fee and ways to pay please contact the Patent Office.

CLONING AND CHARACTERISATION
OF NOVEL MAMMALIAN PEPTIDASES

5 The present invention is concerned with novel
mammalian peptidases and, in particular, with
peptidases designated NAALAD-ase (N-acetylated alpha-
linked acidic dipeptidases), cDNA sequences encoding
for said peptidases and methods of identifying
10 compounds which inhibit or enhance the activity or
expression of such peptidases in addition to the
compounds so identified.

NAALAD-ase I has previously been identified as a
type II membrane glycoprotein with carboxypeptidase
activity and sequence similarity to the transferrin
15 receptor and which may be important in the progression
of prostate cancer, being highly expressed in prostate
tumours. In the CNS it has been postulated to have a
role in modulating neuronal glutamergic activity. The
nucleotide and amino acid sequences for NAALAD-ase I
20 have previously been identified (US 5,538,686).

The present inventors have now surprisingly
identified and cloned cDNA molecules which represent
an expansion of this family of proteins and which
enzymes have never previously been identified or
25 characterised.

Therefore, according to a first aspect of the
present invention, there is provided a cDNA molecule
encoding a peptidase designated human NAALAD-ase II
having the amino acid sequence illustrated in Figure 1
30 or a functional equivalent or derivative or
bioprecursor thereof. Preferably, the cDNA molecule
comprises the sequence of nucleotides illustrated in
Figure 1. Also provided by the present invention are
splice variants of the human NAALAD-ase II protein
35 illustrated in Figure 3 and preferably which splice

variants are encoded by the nucleotide deletions or insertions indicated in Figure 3.

According to a further aspect of the present invention there is provided a cDNA molecule encoding a
5 peptidase designated NAALAD-ase III or IV having the amino acid sequences as illustrated in Figures 4 and 5 respectively. Preferably, the cDNA comprises the sequences of nucleotides illustrated in Figures 4 and 5 respectively. Also encompassed within the present
10 invention are nucleic acid molecules capable of hybridising to the cDNA molecules according to the invention.

Also provided by the present invention are the NAALAD-ase proteins encoded by the DNA sequences
15 according to the invention. Therefore, according to a further aspect of the invention there is provided a human NAALAD-ase II protein having an amino acid sequence encoded by the nucleotide sequence illustrated in Figure 1 or the sequence of Figure 1
20 including the insertions or deletions illustrated in Figure 3. Preferably, the NAALAD-ase II protein comprises a human NAALAD-ase II protein. Also provided by the present invention is a NAALAD-ase III and IV protein having amino acid sequences encoded by
25 the nucleotide sequences illustrated in Figure 4 and 5 respectively or a functional equivalent, derivative or bioprecursor thereof and which protein is preferably of human origin. Preferably, the NAALAD-ase III and IV proteins comprises the amino acid sequences
30 illustrated in Figures 4 and 5 respectively.

The cDNA molecules according to the invention may advantageously be included in an expression vector which may itself be used to transform, transfect or infect a host cell, which cell may be bacterial or
35 eukaryotic in origin. Thus, advantageously, a range

of cells, tissues or organisms may be transfected following incorporation of the appropriate cDNA into the expression vector. The vector may include an appropriate promoter, such as a cytomegalovirus promoter and optionally a sequence encoding a reporter molecule such as, for example, green fluorescent protein.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

The present invention also comprises a transgenic cell, tissue or organism comprising a transgene, capable of expressing a NAALAD-ase protein according to the invention.

The term "transgene capable" of expressing the protein according to the invention, as used herein should be taken to mean a suitable nucleic acid sequence which leads to the expression of said protein according to the invention or a protein having the same function or activity. The transgene may include, for example, genomic nucleic acid isolated from a human source or synthetic nucleic acid, inducing cDNA. The term "transgenic organism, tissue or cell as used herein means any suitable organism and/or part of an

organism, tissue or cell, that contains exogenous nucleic acid either stably integrated in the genome or in an extrachromosomal state. The transgenic cell is preferably a COS cell. Preferably, the transgene
5 comprises an expression vector according to the invention.

The term "functional fragment" as used herein should be taken to mean a fragment of the gene encoding the NAALAD-ase protein according to the
10 invention. For example, the gene may comprise deletions or mutations but may still encode a functional NAALAD-ase protein.

In accordance with the present invention, a defined nucleic acid sequence includes not only the
15 identical nucleic acid but also minor base variations from the natural nucleic acid sequence, including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the same amino acid) due to the degenerate code in
20 conservative amino acid substitution. The term "nucleic acid or cDNA sequence" also includes the complementary sequence to any single stranded sequence given which includes the definition given above regarding base variations.

25 Furthermore, a defined protein, polypeptide or amino acid sequence according to the invention includes not only the identical amino acid sequence but also minor amino acid variations from the natural amino acid sequence including conservative amino acid
30 replacements (a replacement by an amino acid that is related in its side chains). Also included are amino acid sequences which vary from the natural amino acid but result in a polypeptide which is immunologically identical or similar to the polypeptide encoded by the
35 naturally occurring sequence.



The
Patent
Office

GB9 2241

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

REC'D 15 NOV 1999

WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Andrew Gorse

Dated 5 November 1999

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

G 912241
1/77**Request for grant of a patent**

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference SCB/50965/000

2. Patent application number
(The Patent Office will fill in this part)

14 JUL 1998

9815284.6

3. Full name, address and postcode of the or of each applicant (underline all surnames)

JANSSEN PHARMACEUTICA N.V.
Turnhoutseweg 30
B-2340 BEERSE
BELGIUM

Patents ADP number (if you know it)

If the applicant is a corporate body, give the country/state of its incorporation

BELGIUM

531937001

4. Title of the invention

CLONING AND CHARACTERISATION OF NOVEL
MAMMALIAN PEPTIDASES

5. Name of your agent (if you have one)

BOULT WADE TENNANT
27 FURNIVAL STREET
LONDON
EC4A 1PQ

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Patents ADP number (if you know it)

42001

If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number
(if you know it)Date of filing
(day/month/year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

YES

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form -

Description 63

Claim(s) 6

Abstract -

Drawing(s) 16 + 16



10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents (Please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Date

Bouet Wade Tennant

14 July 1998

12. Name and daytime telephone number of person to contact in the United Kingdom **COLM D. MURPHY**
0171 404 5921

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) For details of the fee and ways to pay please contact the Patent Office.

**CLONING AND CHARACTERISATION
OF NOVEL MAMMALIAN PEPTIDASES**

The present invention is concerned with novel
5 mammalian peptidases and, in particular, with
peptidases designated NAALAD-ase (N-acetylated alpha-
linked acidic dipeptidases), cDNA sequences encoding
for said peptidases and methods of identifying
10 compounds which inhibit or enhance the activity or
expression of such peptidases in addition to the
compounds so identified.

NAALAD-ase I has previously been identified as a
type II membrane glycoprotein with carboxypeptidase
activity and sequence similarity to the transferrin
15 receptor and which may be important in the progression
of prostate cancer, being highly expressed in prostate
tumours. In the CNS it has been postulated to have a
role in modulating neuronal glutamergic activity. The
nucleotide and amino acid sequences for NAALAD-ase I
20 have previously been identified (US 5,538,686).

The present inventors have now surprisingly
identified and cloned cDNA molecules which represent
an expansion of this family of proteins and which
enzymes have never previously been identified or
25 characterised.

Therefore, according to a first aspect of the
present invention, there is provided a cDNA molecule
encoding a peptidase designated human NAALAD-ase II
having the amino acid sequence illustrated in Figure 1
30 or a functional equivalent or derivative or
bioprecursor thereof. Preferably, the cDNA molecule
comprises the sequence of nucleotides illustrated in
Figure 1. Also provided by the present invention are
splice variants of the human NAALAD-ase II protein
35 illustrated in Figure 3 and preferably which splice

variants are encoded by the nucleotide deletions or insertions indicated in Figure 3.

According to a further aspect of the present invention there is provided a cDNA molecule encoding a
5 peptidase designated NAALAD-ase III or IV having the amino acid sequences as illustrated in Figures 4 and 5 respectively. Preferably, the cDNA comprises the sequences of nucleotides illustrated in Figures 4 and 5 respectively. Also encompassed within the present
10 invention are nucleic acid molecules capable of hybridising to the cDNA molecules according to the invention.

Also provided by the present invention are the NAALAD-ase proteins encoded by the DNA sequences
15 according to the invention. Therefore, according to a further aspect of the invention there is provided a human NAALAD-ase II protein having an amino acid sequence encoded by the nucleotide sequence illustrated in Figure 1 or the sequence of Figure 1
20 including the insertions or deletions illustrated in Figure 3. Preferably, the NAALAD-ase II protein comprises a human NAALAD-ase II protein. Also provided by the present invention is a NAALAD-ase III and IV protein having amino acid sequences encoded by
25 the nucleotide sequences illustrated in Figure 4 and 5 respectively or a functional equivalent, derivative or bioprecursor thereof and which protein is preferably of human origin. Preferably, the NAALAD-ase III and IV proteins comprises the amino acid sequences
30 illustrated in Figures 4 and 5 respectively.

The cDNA molecules according to the invention may advantageously be included in an expression vector which may itself be used to transform, transfect or infect a host cell, which cell may be bacterial or
35 eukaryotic in origin. Thus, advantageously, a range

of cells, tissues or organisms may be transfected following incorporation of the appropriate cDNA into the expression vector. The vector may include an appropriate promoter, such as a cytomegalovirus promoter and optionally a sequence encoding a reporter molecule such as, for example, green fluorescent protein.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

The present invention also comprises a transgenic cell, tissue or organism comprising a transgene, capable of expressing a NAALAD-ase protein according to the invention.

The term "transgene capable" of expressing the protein according to the invention, as used herein should be taken to mean a suitable nucleic acid sequence which leads to the expression of said protein according to the invention or a protein having the same function or activity. The transgene may include, for example, genomic nucleic acid isolated from a human source or synthetic nucleic acid, inducing cDNA. The term "transgenic organism, tissue or cell as used herein means any suitable organism and/or part of an

organism, tissue or cell, that contains exogenous nucleic acid either stably integrated in the genome or in an extrachromosomal state. The transgenic cell is preferably a COS cell. Preferably, the transgene
5 comprises an expression vector according to the invention.

The term "functional fragment" as used herein should be taken to mean a fragment of the gene encoding the NAALAD-ase protein according to the
10 invention. For example, the gene may comprise deletions or mutations but may still encode a functional NAALAD-ase protein.

In accordance with the present invention, a defined nucleic acid sequence includes not only the
15 identical nucleic acid but also minor base variations from the natural nucleic acid sequence, including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the same amino acid) due to the degenerate code in
20 conservative amino acid substitution. The term "nucleic acid or cDNA sequence" also includes the complementary sequence to any single stranded sequence given which includes the definition given above regarding base variations.

25 Furthermore, a defined protein, polypeptide or amino acid sequence according to the invention includes not only the identical amino acid sequence but also minor amino acid variations from the natural amino acid sequence including conservative amino acid
30 replacements (a replacement by an amino acid that is related in its side chains). Also included are amino acid sequences which vary from the natural amino acid but result in a polypeptide which is immunologically identical or similar to the polypeptide encoded by the
35 naturally occurring sequence.

A further aspect of the invention provides a nucleic acid sequence of at least 15 nucleotides of a nucleic acid according to the invention and preferably from 15 to 50 nucleotides.

5 These sequences may, advantageously be used as probes or primers to initiate replication or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used
10 in diagnostic kits or the like for detecting for the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with a sample under hybridising conditions and detecting for the presence of any duplex formation
15 between the probe and any nucleic acid in the sample. Nucleic acid sequences according to the invention may also be produced using recombinant or synthetic means such as described in Sambrook et al (Molecular Cloning: A Laboratory Manual, 1989). Advantageously,
20 human allelic variants or polymorphisms of the DNA according to the invention may be identified by, for example, probing DNA libraries from a range of individuals for example from different populations. Furthermore, nucleic acids and probes according to the
25 invention may be used to sequence genomic DNA from patients using techniques well known in the art, such as the Sanger Dideoxy chain termination method, which may advantageously ascertain any predisposition of a patient to certain proliferative disorders.

30 A nucleic acid sequence or protein identified according to the present invention may advantageously be used as a medicament, or in the preparation of a medicament for treating cancer or neurodegenerative disorders such as Alzheimer's disease or ALS, and
35 other diseases or disorders mediated by peptidases

according to the present invention. Advantageously, the cDNA and protein according to the invention in addition to nucleic acids hybridisable to the cDNA may also be included in a pharmaceutical composition
5 together with a pharmaceutically acceptable carrier, diluent or excipient therefor. Thus, the composition may comprise antisense nucleic acids for use in, for example, gene therapy.

Further provided by the present invention is a
10 method of determining whether a compound is an inhibitor or an enhancer of NAALAD-ase activity, which method comprises contacting said compound in the presence of [³H]N-acetyl-L-aspartyl-L-glutamate (NAAG), and monitoring for the extent of hydrolysis of
15 said NAAG.

NAALAD-ase peptides have previously been reported to have a role in prostate and other potential cancers in which case compounds identified as either inhibitors or enhancers of NAALAD-ase activity, may
20 advantageously be used as medicaments, or in the preparation of medicaments for treating cancers or neurodegenerative diseases such as Alzheimer's disease or ALS. These compounds may also be included in a pharmaceutical composition together with a
25 pharmaceutically acceptable carrier, diluent or excipient therefor.

A further aspect of the present invention comprises a method of identifying a compound which is an inhibitor or an enhancer of expression of a NAALAD-
30 ase protein according to the invention which method comprises contacting a host cell, tissue or organism expressing said protein with said compound and monitoring the expression of said protein compared to a control which comprises a cell expressing said
35 peptidase according to the invention but which has not

been contacted with said compound.

Preferably, said NAALAD-ase expressing cell comprises a host cell according to the invention or a transgenic cell, tissue or organism as described
5 above. Preferably, said monitoring step comprises monitoring for the expression of said reporter molecule.

Compounds identified according to the above identified method may also be used as a medicament or
10 in the preparation of a medicament for treating neurodegenerative disorders such as Alzheimer's disease, schizophrenia or ALS. Such compounds may also be included in a pharmaceutical composition together with a pharmaceutically acceptable carrier,
15 diluent or excipient therefor to the present invention may also be prepared.

Antibodies to the NAALAD-ase peptidases according to the present invention may, advantageously be prepared according to those techniques known in the
20 art. Polyclonal antibodies against these peptides may be produced by immunizing animals using the selected peptides. Monoclonal antibodies are prepared using hybridoma technology by fusing antibody producing B cells from immunized animals with myeloma cells and
25 selecting the resulting hybridoma cell line producing the desired antibody. Alternatively, monoclonal antibodies may be produced by *in vitro* techniques known to a person of ordinary skill in the art.

The antibodies prepared may be used in a method
30 for detecting for the presence of said NAALAD-ase peptidases according to the invention by reacting said antibodies with a sample to be tested and identifying any protein bound thereto. A kit for such a method may also be provided and which comprises an antibody
35 according to the aspect of the invention described

above and means for reacting said antibody with said sample.

The present invention may be more clearly understood by the following examples, which are purely exemplary, with reference to the accompanying drawings wherein:

Figure 1: nucleotide sequence and amino acid sequence of human NAALAD-ase II. The nucleotide and predicted single letter code amino acid sequence are shown. The putative membrane spanning domain, deduced from hydrophilicity plots, is marked by a line. Potential N-glycosylation sites are shaded.

Figure 2: is an alignment of the predicted protein sequences for human and rat NAALAD-ase II. The amino acid sequences were aligned using the ClustalW alignment program (EMBL, Heidelberg, Germany). Amino acid residues identical in both proteins are highlighted in black. Amino acid residues are numbered in the right hand margin.

Figure 3: is an illustration of alternative splicing of NAALAD-ase II. Amino acid sequence for NAALAD-ase II is shown. Sites at which putative DNA sequences are spliced out are marked by an arrow with the resulting (in-frame) amino acid deletions highlighted in bold italicised letters. Sites of putative intronic DNA insertion are marked by triangles, with the intronic DNA sequence shown above. Resulting changes to the amino acid sequence are highlighted in bold italicised letters. Numbering of amino acid residues is the right.

Figure 4: is a nucleotide sequence and amino acid

sequence of human NAALAD-ase III. The nucleotide and predicted one letter code amino acid sequence are shown. The putative membrane spanning domain, deduced from hydrophilicity plots, is marked by a line.

5 Potential N-glycosylation sites are shaded.

Figure 5: is a nucleotide sequence and amino acid sequence of human NAALAD-ase IV. The nucleotide and predicted one letter code amino acid sequence are shown. The putative membrane spanning domain, deduced from hydrophilicity plots, is marked by a line.

10 Potential N-glycosylation sites are shaded.

Figure 6: is an alignment of the predicted protein sequences for human NAALAD-ases I, II, III & IV. The amino acid sequences were aligned using the ClustalW alignment program. Amino acid residues identical to all four proteins are shaded in black. Amino acid residues identical to three of the four proteins are shaded in grey. Amino acid residues are numbered to the right. A putative Zn^{2+} peptidase domain is highlighted between arrows and was identified by comparison to yeast and bacterial aminopeptidases. Putative residues involved the catalytic site of the α/β hydrolase fold family of proteins are marked by three arrows (nucleophile-acid-base).

15 Figure 7: is a phylogram of NAALAD-ase I, II, III and IV. The human and rat sequences were used and the alignments performed with the CLUSTALW program. The tree was constructed using the GCG 'Distances' program with standard parameters and the 'Growtree' program with the UPGMA method.

30 Figure 8: is an alignment of the NAALAD-ase peptidase

domains with related peptidases.

Amino acid sequences were aligned using the standard settings of ClustalW alignment program. Similar amino acid residues conserved in proteins are shaded in black. Similar amino acid residues conserved in 80% of the proteins are shaded in dark grey. Similar amino acid residues conserved in 60-79% of the proteins are shaded in light grey. Amino acid residues are numbered to the right. Putative residues involved in zinc binding are marked by asterisks. The base residue thought to be important in catalysis is marked by an arrow. Sequence names other than NAALAD-ases correspond to sequence accession numbers in Swiss-Prot and SPTREMBL; Ape 3 yeast, *Saccharomyces cerevisiae* aminopeptidase Y; P96152, *Vibrio cholerae* aminopeptidase; .Ampx vibpr, *Aeromonas proteolytica* aminopeptidase, Apx strgr, *Streptomyces griseus* aminopeptidase. Putative residues involved in zinc binding are marked by asterisks. General base residue thought to be important in catalysis is marked by an arrow.

Figure 9: is an illustration of time course of NAALAD-ase activity. Activity of membrane preparations from COS cells transiently transfected with either NAALAD-ase I (A) or NAALAD-ase III (B). Hydrolysis of 500 nM 3[H]-NAAG in 1mM ZnCl₂, 50 mM Tris.HCL (pH 7.4) was assayed for 0, 15, 30 and 60 min at 37C in the presence (open circles) or absence (closed circles) of 30 μ M quisqualate. Reactions were terminated with 250 mM ice cold sodium phosphate and released 3[H]-glutamate measured. (C) Inhibition of NAALAD-ase I (open circles) and NAALAD-ase III (closed circles) activity by increasing concentrations of quisqualic acid expressed as % of control (activity in the

absence of any inhibitor).

Figure 10: is an illustration of NAALAD-ase activity determinations using SPA beads . (A) 10 μ g of protein from LNCaP cell membrane preparations was used per reaction to develop a high throughput assay NAALAD-ase assay. Hydrolysis of 40 nM 3[H]-NAAG in 1mM ZnCl₂, 50 mM Tris-HCl (pH 7.4) was measured at 25C for different incubation times in a 100 μ l reaction volume. Reactions were terminated by the addition of 100 μ l glycine buffer (pH 3.0) and naked YS-SPA beads. (C) Inhibition of NAALAD-ase I (open circles) or LNCaP membrane NAALAD-ase (closed circles) activity by increasing concentrations of quisqualic acid expressed as a % of control (activity in the absence of any inhibitor).

Figure 11: is an illustration of chromosomal localisation of human NAALAD-ases. Diagrams of FISH mapping results for NAALAD-ase I (A), II (C), III (E) and IV (G). In each diagram dots represent the double FISH signals detected on identified chromosomes. Examples of FISH mapping of NAALAD-ase I (B), II (D), III (F) and IV (H). Left panels show the FISH signals on the identified chromosome marked by a white arrow, right panels show the same mitotic Figure stained with DAPI to identify the chromosome.

Figure 12: is an illustration of northern blot analysis of NAALAD-ase III and IV expression. Human multiple tissue Northern blots (Clontech) containing 2 μ g per lane of poly (A)+ RNA were hybridised with a ³²P-labelled NAALAD-ase III (A) or IV (B) probes at 68°C and washed at high stringency (final washes at 50°C with 1/10 dilution of SSC/0.1 % SDS).

Autoradiographic exposure was done after 3 to 7 days at -70°C with two intensifying screens. Molecular mass markers are indicated in bp on the left hand side of each autoradiogram.

5

Figure 13: is an illustration of RT-PCR analysis of NAALAD-ase expression in different tissues. PCR amplifications with primers specific for human NAALAD-ase I (A), II (B), III (C) and IV (D) or GAPDH (E) were performed on normalised human MTC cDNATM for 25, 30 and 35 cycles. Images from ethidium bromide stained 1.5 % agarose gels were captured using an EagleEye system and EagleSight software (Stratagene). Positive control amplifications were performed using the appropriate NAALAD-ase DNA construct. Negative control amplifications contained reaction mix, enzyme and no DNA template. Positive and negative control reactions were performed for 35 cycles. Arrows highlight a specific size in bp as determined using a 100 bp ladder.

10

15

20

Figure 14: is an illustration of RT-PCR analysis of NAALAD-ase expression in brain areas. PCR amplifications for 25, 30 and 35 cycles, with primers specific for human NAALAD-ase I (A), II (B), III (C) and IV (D) or GAPDH (E) were performed on normalised human cDNAs prepared from dissected brain areas. Images from ethidium bromide stained 1.5 % agarose gels were captured using an EagleEye system and EagleSight software (Stratagene). Results from 30 and 35 cycles of amplification are shown. Positive control amplifications were performed using the appropriate NAALAD-ase DNA construct. Negative control amplifications contained reaction mix and enzyme with no DNA template. All control reactions were performed

25

30

35

for 35 cycles. Arrows highlight a specific size in bp as determined from the 100 bp ladder.

Figure 15: is an illustration of RT-PCR analysis of
5 NAALAD-ase expression in prostate tumour cells. PCR
amplifications with primers specific for human NAALAD-
ase I (A), II (B), III (C) and IV (D) or GAPDH (E)
were performed on normalised cDNAs prepared from cell
lines derived from prostate tumours or on prostate
10 tumour. Images from ethidium bromide stained 1.5 %
agarose gels were captured using an EagleEye system
and EagleSight software (Stratagene). Results from 25,
30 and 35 cycles of amplification are shown. Positive
control amplifications were performed using the
15 appropriate NAALAD-ase DNA construct. Negative control
amplifications contained reaction mix and enzyme with
no DNA template. All control reactions were performed
for 35 cycles. Table 1.

20 Native Taq polymerase and PCR buffer with MgCl₂, the
ExpandTM long template PCR system, ampicillin, IPTG
(isopropyl- β -D-thiogalactoside), X-gal (5-bromo-4-
chloro-3-indolyl- β -D-galactopyranoside) and all
restriction enzymes used were from Boehringer Mannheim
25 (Mannheim, Germany). Super Taq polymerase was from HT
Biotechnology (Cambridge, UK). 10 mM dNTP mix was
purchased from Life Technologies (Gaithersburg, MD,
USA). The original TA cloning kit and the expression
vector pcDNA-3 were purchased from Invitrogen BV
30 (Leek, The Netherlands). The Qiagen plasmid mini- and
maxi- DNA purification kits, the Qiaquick gel
extraction kit and the Qiaquick PCR purification kit
were purchased from Qiagen GmbH (Düsseldorf, Germany).
MarathonTM Ready cDNA kits, MTN cDNA panels and MTN
35 Northern blots were obtained from Clontech

Laboratories (Palo Alto, CA, USA). The QuickChange™ Site-Directed Mutagenesis Kit was purchased from Stratagene GmbH (Heidelberg, Germany). All PCR reactions were performed in a GeneAmp PCR system 9600
5 cyclor (Perkin Elmer, Foster City, CA, USA). LB (Luria-Bertani) medium consists of 10 g/l of tryptone, 5 g/l of yeast extract and 10 g/l of NaCl. 2x YT/ampicillin plates consist of 16 g/l of tryptone, 10 g/l of yeast extract, 5 g/l of NaCl, 15 g/l of agar
10 and 100 mg/l of ampicillin.

Expressed sequence tag clones.

Clone numbers 4190746, 1547649, 3448872, 3608639, 2615389 and 1333965 were ordered from the LifeSeq™
15 expressed sequence tag (EST) database (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA). Samples were delivered as transformed bacterial clones.

Oligonucleotide synthesis for PCR and DNA sequencing.
20 All oligonucleotide primers were synthesised by the β -cyanoethylphosphoramidite chemical method on a PerSeptive Biosystems (Framingham, MA, USA) Expedite MOSS Synthesiser or ordered from Eurogentec (Seraing, Belgium). Insert-specific sequencing primers (15- and
25 16-mers) were designed manually. DNA was prepared on Qiagen-tip-20 or -100 anion exchange or Qiaquick spin columns (Qiagen GmbH, Düsseldorf, Germany) and recovered from the columns in 30 μ l TE-buffer (10 mM Tris.HCL, 1 mM EDTA (sodium salt), pH 8.0).
30 Sequencing reactions were done on both strands using the ABI prism BigDye Terminator Cycle sequencing kit and were run on an Applied Biosystems 377XL sequencer (Perkin Elmer, ABI Division, Foster City, CA, USA). The Sequencher™ software was used for sequence
35 assembly and manual editing (GeneCodes, Ann Arbor, MI,

USA).

Cloning and sequence analysis of NAALAD-ase I, II, III and IV.

- 5 Sequence similarity searching for NAALAD-ase like molecules. Using the complete human (Accession no. M99487), rat (Accession no. RNU75973) and mouse (Accession no. AF026380) NAALAD-ase I protein sequences, the complete rat NAALAD-ase II (Accession
- 10 no. AF009921) protein sequence and a partial human NAALAD-ase II protein sequence (Accession no. AF010141) as query sequences, a BLAST (Basic Local Alignment Search Tool; Altschul et al., 1990) search was performed on the WashU Merck expressed sequence
- 15 tag (EST) database and on a proprietary LifeSeqTM human EST database (Incyte Pharmaceuticals Inc, Palo Alto, CA, USA). Six EST clones with homology to NAALAD-ase I and II were ordered from the Incyte Pharmaceuticals.
- 20 Plasmid DNA preparation and sequencing from Incyte clones. Each Incyte bacterial clone was grown overnight at 37°C in 100 ml LB-medium supplemented with 100 µg/ml of ampicillin. Plasmid DNA was prepared
- 25 using a Qiagen plasmid midi or maxi kit according to the manufacturer's instructions. The DNA inserts of all clones were completely sequenced on both strands. Clone 4190746 contained sequences corresponding to human NAALAD-ase II and also included likely intronic
- 30 sequences, clones 133965, 1547649, 3448872, 3608639 contained overlapping DNA sequences coding for a novel protein with some similarity to other NAALAD-ases and clone 2615389 contained DNA sequences corresponding to yet another novel protein with some similarity to
- 35 previously identified NAALAD-ases. The novel protein
-

sequences derived from these clones were termed
NAALAD-ase III and NAALAD-ase IV respectively.

Cloning of NAALAD-ase I by PCR. Sequence data from
5 human NAALAD-ase I (Accession no. M99487) was used to
design primers to amplify the complete coding sequence
of NAALAD-ase I by PCR. Primers used were NAALD1S2
(BamHI) = 5'- CCC GGA TCC GAG ATG TGG ATT CTC CTT CAC
GAA AC -3' and NAALD1AS2(XhoI) = 5'- CCC CTC GAG TTA
10 GGC TAC TTC ACT CAA AGT CTC TGC -3' (restriction sites
to be introduced are underlined). PCR amplification
was performed using Human Marathon-Ready™ cDNA from
prostate in a total reaction volume of 50 µl,
containing 1X Expand Long Template™ PCR buffer 2, 0.2
15 mM of each dNTP, 0.2 µM each of oligonucleotide
primers NAALD1S1(BamHI) and NAALD1AS1(XhoI), 1 µl of
Marathon-Ready™ cDNA and 2.5 U of Expand Long
Template PCR mix. Samples were pre-heated at 94°C for
5 min before addition of enzyme. Cycling was for 45 s
20 at 94°C, 1 min at 55°C and 1 min 48 s at 68°C for 30
cycles, with a final step of 7 min at 72°C. PCR
products were analysed on a 1% agarose gel (wt / vol)
in 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA
(sodium salt), pH 8.3) and the most prominent DNA band
25 was excised from the gel and purified with the
Qiaquick gel extraction kit (Qiagen GmbH, Düsseldorf,
Germany). The resulting 2303 base pair (bp) fragment
was cloned into the plasmid vector pCR2.1 using the
original TA cloning kit according to the
30 manufacturer's instructions. Approximately 20 ng of
purified fragment was ligated to 50 ng of pCR2.1
plasmid DNA with 4 U T4 DNA ligase in a total volume
of 10 µl. Ligations were incubated overnight at 14°C.
2 µl of the ligation reaction was transformed into
35 TOP10F' competent cells using heat-shock

transformation and plated on 2x YT / ampicillin plates supplemented with IPTG and X-gal for blue-white screening. Colony screening was performed on 10 white colonies, from which plasmid DNA was prepared using the Qiagen plasmid mini DNA purification kit, and then digested with BamHI and XhoI . Four clones containing an insert of the appropriate size were sequenced fully. All the clones had at least one mis-sense PCR error. Cl. 10.0 with a single PCR error at position 1183 was used as the template for a site directed mutagenesis (SDM) reaction using the QuickChange™ SDM Kit. Reactions were carried out according to the manufacturer's instructions. Primers designed for the amplification reactions were NAALD1-SDM-S1 = 5'- CCC TCA GAG TGG AGC AGC TGT TGT TCA TGA AAT TGT GAG G -3' and NAALD1-SDM-AS1 = 5'- CCT CAC AAT TTC ATG AAC AAC AGC TGC TCC ACT CTG AGG G -3'. Three white clones from the SDM transformations were screened. Plasmid DNA was prepared using the Qiagen plasmid mini DNA purification kit, digested with BamHI and XhoI and sequenced around the mutation site . A single clone (cl. 2.0) was sequenced fully on both strands to confirm the complete correct NAALAD-ase I sequence.

Cloning of NAALAD-ase II by PCR and 5' rapid amplification of cDNA ends (RACE) PCR. Sequence data from partial human NAALAD-ase II (Accession no. AF10141) was used to design primers to amplify the 3' end of NAALAD-ase II by PCR. Primers used were NAALD2S1 = 5'- GTT CTT CAA CAA GCT GCA GGA GCG -3' and NAALD2AS1(XhoI) = 5'- CCC CTC GAG CCG GAG TAA AGG GAG GGC TGA AG -3'. Human Marathon-Ready™ cDNA from brain, fetal brain, prostate, small intestine, colon were used in the amplification reactions. First round PCR amplification was performed in a total reaction

volume of 50 μ l containing 1X Expand High FidelityTM PCR buffer, 0.2 mM of each dNTP, 0.2 μ M each of oligonucleotide primers NAALD2S1 and NAALD2AS1, 1 μ l of Marathon-ReadyTM cDNA and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 45 s at 94°C, 1 min at 58°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1% agarose gel in 1X TAE buffer. Second round PCR amplification was performed with nested primers NAALD2S2 = 5'- GGC GAC CTG AGC ATC TAC GAC AAC -3' and NAALD2AS2 (XhoI) = 5'- CCC CTC GAG TCC CCT CAG AGG TCA GCC ACA G -3'. 1 μ l of the first round amplification reaction in a total volume of 50 μ l containing 1X Expand High FidelityTM PCR buffer, 0.2 mM of each dNTP, 0.2 μ M each of oligonucleotide primers NAALD2S2 and NAALD2AS2 and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 45 s at 94°C, 1 min at 57°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1 % agarose gel in 1X TAE buffer and the most prominent DNA bands were excised from the gel and purified with the Qiaquick gel extraction kit. The resulting fragments 600-800 bp were cloned into pCR2.1 as described previously. Ligations were incubated overnight at 14°C and transformed into TOP10F' competent cells and plated on 2x YT / ampicillin plates supplemented with IPTG and X-gal for blue-white screening. Colony screening was performed on five white colonies from each transformation. Plasmid DNA was prepared from these colonies using the Qiagen plasmid mini DNA purification kit and then digested with EcoRI . Plasmids containing inserts of the

appropriate size were end sequenced using vector primers and then the full sequencing on both strands of putative NAALAD-ase II clones was performed. Clone 6.9 derived from small intestine extended to the translation termination codon.

To obtain unknown 5' coding sequence for human NAALAD-ase II, two anti-sense primers were designed for 5' rapid amplification of cDNA ends (5' RACE). The primers were NAALD2AS3 = 5'- GCC AGC ACC CAG AGA ACC CAA G -3' and NAALD2AS4 = 5'- GCT GCG GTT GAA GTA CCG GAT C -3'. Human Marathon-Ready™ cDNA from brain, fetal brain, prostate, small intestine and colon were used for the 5' RACE according to the manufacturer's instructions. The Marathon-Ready™ cDNA was prepared using oligo-dT priming and a Marathon cDNA adaptor (including two different adaptor-primer annealing sites) ligated to the 5' end of the cDNA. Adaptor-primer AP1 (5'- CCA TCC TAA TAC GAC TCA CTA TAG GGC - 3') and nested adaptor-primer AP2 (5'- ACT CAC TAT AGG GCT CGA GCG GC -3') were included in the kit. First round PCR amplification was performed in a total reaction volume of 50 µl containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 µM each of oligonucleotide primers AP1 and NAALD2AS3, 5 µl of Marathon-Ready™ cDNA and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 58°C and 2 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1% agarose gel in 1X TAE buffer. Second round 5' RACE was performed using 1 µl of the first round amplification reaction in a total volume of 50 µl containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 µM each of oligonucleotide primers AP2 and NAALD2AS4 and 2.5 U of Expand High

Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 57°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were excised from the gel and cloned into the vector pCR2.1 as described earlier. Colony screening by PCR was performed on 60 white colonies in 45 µl PCR mix containing 1X PCR buffer with MgCl₂, 0.2 mM dNTP, 0.5 µM each of vector primer M13FOR (5'- TGT AAA ACG ACG GCC AGT -3') and M13REV (5'- CAG GAA ACA GCT ATG ACC -3') and 0.35 U of super Taq DNA polymerase. Colonies were picked from the plates, inoculated into 100 µl LB medium supplemented with 100 µg/ml of ampicillin and incubated for 1 hr at 37°C. 5 µl of the incubated culture was then added to 45 µl PCR mix. PCR was performed for 30 cycles (45 s at 95°C, 1 min at 48°C and 50 s at 72°C). 20 µl of the PCR reactions was analysed on 1 % agarose gel in 1x TAE buffer. Clones containing inserts were grown overnight in 3 ml LB medium supplemented with 100 µg/ml of ampicillin and plasmid DNA was prepared using the Qiagen plasmid mini DNA purification kit. From the 51 clones sequenced one clone 4.10 contained 258 bp of sequence 5' to NAALD2AS4 of which 70 bp were novel. The 5' RACE clone (cl. 4.10) and 3' PCR clone (cl. 6.9) were both digested with BamHI. The digested material was run on a 1 % agarose gel in 1X TAE buffer. Two bands were excised, a 336 bp band from cl. 4.10 containing all additional 5' RACE DNA sequences and a ~4700 bp fragment containing the remaining 3' NAALAD-ase II and vector sequence from cl. 6.9. The gel slices were purified with the Qiaquick gel extraction kit. The larger of the two fragments was dephosphorylated with 1.5 U calf of intestinal alkaline phosphatase for 1 hr at 37C and then heat

inactivated for 20 min at 75°C, in order to prevent the self ligation of the fragment with itself. Ligations were performed as described previously and 2 µl of the reaction mixture transformed into 35 µl DH10b

5 electrocompetent cells by electroporation (single pulse; 2500V, 25µF, 201 W, 5 ms). The electroporated sample was added to 1 ml SOC media and incubated for 1 hr at 37°C before 100 µl of the culture was plated on to 2x YT / ampicillin plates. Colonies were picked the

10 following day, plasmid DNA prepared and tested by restriction digest. A single clone (cl. 2.0) was fully sequenced on both strands and found to contain the complete 3' coding sequence and the additional sequence from the first 5' RACE reactions.

15 To obtain additional 5' coding sequence for the human NAALAD-ase II, two new anti-sense primers were synthesised corresponding to sequences from Incyte clone number 4190746. Primers used were NAALD2AS5 = 5'- CTG CAG CTT GTT GAA CTC TTC TGT G -3' and

20 NAALD2AS6 = 5'- CAA ACA CGA TTG ATC TGC GAG GAC -3'. Human Marathon-Ready™ cDNA from brain, fetal brain, prostate, small intestine, colon and heart were used for the 5' rapid amplification of cDNA ends (5' RACE) according to the manufacturer's instructions. First

25 round PCR amplification was performed in a total reaction volume of 50 µl containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 µM each of oligonucleotide primers AP1 and NAALD2AS5, 5 µl of Marathon-Ready™ cDNA and 2.5 U of Expand High Fidelity

30 PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 58°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1% agarose gel in 1X TAE buffer (40 mM

35 Tris-acetate, 1 mM EDTA, pH 8.3). Second round PCR

amplification was performed using 1 μ l of the first round amplification reaction in a total volume of 50 μ l containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 μ M each of oligonucleotide primers AP2 and NAALD2AS6 and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 57°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1 % agarose gel in 1X TAE buffer and the 8 most prominent DNA bands were excised from the gel and purified with the Qiaquick gel extraction kit. The resulting fragments (400-1300bp) were cloned into the plasmid vector pCR2.1 using the original TA cloning kit. Ligations and transformations were performed as described earlier before plating on 2x YT / ampicillin plates supplemented with IPTG and X-gal for blue-white screening. Colony screening was performed on five white colonies from each transformation. Plasmid DNA was prepared from these colonies using the Qiagen plasmid mini DNA kit and digested with EcoRI . Plasmids containing inserts of the appropriate size were end sequenced using vector primers and then the full sequence of putative NAALAD-ase II clones was determined using primer walking. The DNA sequence of five clones, from a small intestine cDNA, extended the coding sequence for NAALAD-ase II in the 5' direction beyond the putative translation start codon and included part of the 5' untranslated region. One of the these clones one (cl. 2.2) was used for further experiments. To construct a full length NAALAD-ase II clone, two new primer sets were designed to introduce a unique restriction site (Mun I) into the DNA sequence of NAALAD-ase II without resulting in a change in amino

acid sequence or frame shift in the open reading frame (ORF). The first primer set was NAALD2S3 (EcoRV) = 5'-CGG ATA TCC GCA GGA TGC AGT GGA CGA AG -3' and NAALD2AS8 (MunI) = 5'- CAA ACA CAA TTG ATC TGC GAG GAC GC -3' and the second primer set was NAALD2S8 (MunI) = 5'- GCG TCC TCG CAG ATC AAT TGT GTT TG -3' and NAALD2AS1 (XhoI). PCR amplification was performed on 1 μ l cl. 2.0 plasmid DNA with primers NAALD2S3 (EcoRV) and NAALD2AS8 (MunI) or on 1 μ l cl. 2.2 plasmid DNA with primers NAALD2S8 (MunI) and NAALD2AS1 (XhoI). Total reaction volumes were as previously described. Samples were preheated at 94°C for 5 min before addition of the Expand High Fidelity enzyme. Cycling was for 45 s at 95°C, 1 min at 57°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1% agarose gel in 1X TAE buffer and TA cloned into pCR2.1 as previously described.

Splice variant analysis of NAALAD-ase II.
In PCR reactions using NAALAD-ase II specific primers a number of amplified PCR products were observed of unexpected size. These PCR reactions were repeated using a number of different cDNAs and bands were excised and purified prior to original TA cloning, as described previously. Plasmid DNA was prepared from these clones and inserts were fully sequenced on both strands in order to identify possible splice variants.

Cloning of NAALAD-ase III by PCR and 5' rapid amplification of cDNA ends (RACE) PCR.
Sequencing results from Incyte clone 3608639 suggested that this clone contained DNA sequence spanning the complete coding sequence, 2220 bp in size, of a putative NAALAD-ase like molecule (NAALAD-ase III)

that had similar sequence to NAALAD-ase I and II. To confirm that there was no possible initiation codon upstream of the initiation codon already determined 5' RACE PCR was performed. Two anti-sense primers were
5 designed for 5' RACE based on the sequence derived from the clone 3608639, NAALD3AS1 = 5'- CTT TGA TGA TAG CGC ACA GAA GTG G -3' and NAALD3AS2 = 5'- GGA AAG ATG CCA GCG CAG GAC -3'. Human Marathon-Ready™ cDNA from brain, foetal brain, prostate, small intestine
10 and colon were used for the 5' RACE according to the manufacturer's instructions. First round PCR amplification was performed in a total reaction volume of 50 µl containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 µM each of
15 oligonucleotide primers AP1 and NAALD3AS1, 5 µl of Marathon-Ready™ cDNA and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 58°C and 2 min at 72°C for 30 cycles,
20 with a final step of 7 min at 72°C. Second round 5' RACE was performed using 1 µl of the first round amplification reaction in a total volume of 50 µl containing 1X Expand High Fidelity™ PCR buffer, 0.2 mM of each dNTP, 0.2 µM each of oligonucleotide primers
25 AP2 and NAALD3AS2 and 2.5 U of Expand High Fidelity PCR mix. Samples were preheated at 94°C for 5 min before addition of the enzyme. Cycling was for 30 s at 94°C, 30 s at 57°C and 1 min at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were
30 analysed on a 1 % agarose gel in 1X TAE buffer and the most prominent DNA bands were excised from the gel and purified with the Qiaquick gel extraction kit. The resulting fragments (250 - 600 bp) were cloned into the plasmid pCR2.1 as described previously. 32 white
35 colonies were grown overnight in 3 ml LB medium

supplemented with 100 μ g/ml of ampicillin and plasmid DNA was prepared using the Qiagen plasmid mini DNA purification kit. No upstream initiation codon was identified from any of the 32 clones analysed.

5

Cloning of NAALAD-ase IV by PCR.

Sequencing results from Incyte clone 2615389 revealed that this clone contained partial coding sequence and 3' UTR, of another putative NAALAD-ase like molecule (NAALAD-ase IV) that was related in sequence to NAALAD-ase I, II and III. The DNA sequence obtained was used in a BLAST search on the Incyte LifeSeqTM EST database. One contig (2519841) was assembled from 150 overlapping Incyte EST sequences that spanned 1881 bp and contained a coding region of 1419 bp. The sequence data from human NAALAD-ase IV contig 2519841 was used to design primers to amplify the complete coding sequence of by PCR. Primers used were NAALD4SP2 = 5'-CGT CAG AGC CGC CCT ATC AGA TTA TC -3' and NAALD4AP4 = 5'-GAG GAG TTT TCC AAA GTT GCA GAC CC -3'. PCR amplification was performed using a human hippocampal cDNA in a total reaction volume of 50 μ l, containing 1X Expand High FidelityTM PCR buffer, 0.2 mM of each dNTP, 0.2 μ M each of oligonucleotide primers NAALD4SP2 and NAALD4AP4, 1 μ l of cDNA and 2.5 U of Expand High FidelityTM PCR mix. Samples were pre-heated at 95°C for 5 min before addition of enzyme. Cycling was for 45 s at 94°C, 1 min at 58°C and 35 s at 72°C for 30 cycles, with a final step of 7 min at 72°C. PCR products were analysed on a 1% agarose gel in 1X TAE buffer and the most prominent DNA band was excised from the gel and purified with the Qiaquick gel extraction kit. The resulting 1544 bp DNA fragment was sub-cloned into the plasmid vector pCR2.1-TOPO using the TA TOPO cloning kit according to the

manufacturer's instructions. Approximately 10 ng of purified fragment was ligated to 10 ng of pCR2.1-TOPO plasmid DNA. Ligations were incubated for 5 min at 25°C. Transformations into TOP10F' competent cells and colony screening was performed as previously described. Three clones containing an insert of the correct size were sequenced fully and two clones were found to contain no PCR errors (cl. 28.0 and cl.1.0).

10 Activity determinations of NAALAD-ases transiently expressed in COS cells.

Sub-Cloning of NAALAD-ases into expression vectors. NAALAD-ase I, II, III and IV clones were each sub-cloned into the cytomegalovirus (CMV) promoter-based plasmid vector pCDNA-3. NAALAD-ase I/pCR2.1 (cl.2.0) was digested with BamHI/XhoI to excise the complete NAALAD-ase I sequence. NAALAD-ase II-5'/pCR2.1 (cl.2.0) was digested with EcoRV/MunI and NAALAD-ase II-3'/pCR2.1 (cl.2.2) was digested with MunI/XhoI to excise the two halves of complete NAALAD-ase II sequence. NAALAD-ase III/pINCYTE was digested with EcoRI to excise the complete NAALAD-ase III sequence. NAALAD-ase IV/pCR2.1 (cl.28.0) was digested with HindIII/XbaI to excise the complete NAALAD-ase IV sequence. Fragments were gel-purified and sub-cloned into dephosphorylated pCDNA3 that had been previously digested with the appropriate restriction enzymes. The resulting expression constructs were verified by complete sequence analysis as described earlier.

30 Transient transfection into COS cells.

COS cells were maintained in complete medium (DMEM supplemented with 10% foetal calf serum, 1X non-essential amino acids and a 1X streptomycin/penicillin/glutamine mix). Prior to

transfection cells were washed twice with PBS, Ca
2+/Mg 2+ free pre-warmed to 37°C and dissociated from
culture plates by addition of a trypsin/EDTA solution
for 1-2 minutes. Cells were collected after a 10 min
5 spin at 1600 rpm and re-suspended in 1-2 ml of
equilibrated (5% CO₂/95% air) pre-warmed medium. The
cell titre was determined in a Coulter counter and
cells were seeded in 6 well plates at a density of
15000 cells/cm² and allowed to reach approximately 80%
10 confluence (~2 days growth).

For each transfection 6 µl FuGENE6 (Boehringer
Mannheim, Germany) was added to 96µl serum free
medium and incubated for 5 minutes at 20°C. This
preparation was added to a second tube containing 1 µg
15 of NAALAD-ase/pcDNA3 DNA, mixed gently and allowed to
stand for 15 minutes at RT. The DNA/FuGENE6/serum-free
medium mix was pipetted into a well containing 2 ml
fresh complete medium. Cells were incubated for 72 hrs
in a 37° C incubator before collection.

20

Determination of Biological Activity of NAALAD-ase
homologues.

NAALAD-ase activity was quantified by determination of
the extent of hydrolysis of N-acetyl-L-aspartyl-L-
25 (3,4-[³H])-glutamate, as previously described (Blakely
et al., 1988). Transfected COS cell pellets were
incubated with 50 mM Tris-HCl (pH 7.4) / 0.1% Triton
X-100 and vortexed. Homogenates were put through at
least one freeze/thaw cycle in liquid N₂ before being
30 used for assay. All assays were performed in a total
volume of 200 µl containing 50 mM Tris-HCl (pH 7.4), 1
mM Zn²⁺Cl₂, 500 nM N-acetyl-L-aspartyl-L-(3,4-3[H])-
glutamate ([³H]-NAAG) and 10-200 µg membrane protein
in the presence or absence of 30 µM quisqualate / 1%
35 DMSO (vol /vol). Assays were initiated by the addition

of the membrane homogenates to the assay mixture, preincubated at 37°C. The assay mixture was vortexed and incubated at 37°C for various time points. Reactions were terminated by addition of 200 μ l ice cold 250 mM potassium phosphate and placing the reaction tubes in ice water. After a 5 min centrifugation step at 14000 rpm, the assay samples were loaded onto 4 cm anion exchange mini-columns (Bio-Rad AG1-X 8) that had been pre-washed with MilliQ water. Columns were rinsed twice with 2 ml MilliQ water, and then twice with 2 ml of 0.5 M formic acid to selectively elute the [3H]-glutamate. The remaining substrate could be eluted with 7.5M formic acid. Assay eluates were diluted with scintillation cocktail (16 ml Ultima Gold XR) and counted in a scintillation counter (Packard). Blanks (not exposed to membranes) were always subtracted. Inhibition curves were performed under similar conditions with concentrations of quisqualate (QA) ranging up to 300 μ M. Reactions were followed for 30 min for NAALAD-ase I and for 60 min for NAALAD-ase II and III prior to termination of the reaction.

SPA development for high throughput screening. NAALAD-ase assays were performed at 37°C or 25°C for 10 min in the presence of 50 mM Tris-HCl, pH 7.4, containing 1 mM ZnCl₂. Each assay was performed in a 100 μ l reaction volume containing 40 nM 3[H] -NAAG, 50 mM Tris-HCl (pH 7.4), 1 mM Zn²⁺-Cl₂ and either 10 μ g of recombinant NAALAD-ase I or LNCaP cell membrane preparations. Reactions were terminated by addition of 100 μ l glycine buffer (pH 3.0) and 1 mg naked YS-SPA beads. After allowing the beads to settle for 30 min, the assay mixture was counted in a microtiterplate scintillation counter.

Chromosomal localisation of NAALAD-ases by FISH.
Chromosomal mapping studies were carried out by SeedNA
Biotech Inc. (North York, Ontario, Canada) using
fluorescent *in situ* hybridisation (FISH) analysis.

- 5
- Slide preparation. Lymphocytes isolated from human
blood were cultured in α -minimal essential medium
(MEM) supplemented with 10% foetal calf serum and
phytohaemagglutinin at 37 C for 68-72 h. The
- 10 lymphocyte cultures were treated with BrdU (0.18
mg/ml, Sigma Chemical Company, St Louis, MO, USA) to
synchronise the cell population. The synchronised
cells were washed three times with serum-free medium
to release the block and re-cultured at 37C for 6 h in
- 15 α -MEM with thymidine (2.5 μ g/ml, Sigma). Cells were
harvested and slides were prepared using standard
procedures including hypotonic treatment, fixation and
air-drying.
- 20 FISH detection. Probes (NAALAD-aseI/PCR2.1 (cl. 2.0)
cut with BamHI/XhoI; NAALAD-aseII/PCR2.1 (cl. 2.0) cut
with XhoI; NAALAD-aseIII/pcDNA3 (cl. 3.0) cut with
HindIII/XbaI; NAALAD-aseIV/PCR2.1 (cl. 28.0) cut with
HindIII/XbaI) were biotinylated with dATP for 1 h at
- 25 150C using the BioNick labelling system (Life
Technologies, Gaithersburg, MD, USA). The procedure
for FISH detection was performed as previously
described (Heng et al, 1992; Heng & Tsui, 1993).
Slides were baked at 550C for 1 h. After RNase
- 30 treatment, the slides were denatured in 70% formamide
in 2x SSC (20x SSC being 3 M NaCl, 0.3 M sodium
citrate, pH 7.0) for 2 min at 700C followed by
dehydration with ethanol. Probes were denatured at 75
C for 5 min in a hybridisation mix consisting of 50%
- 35 formamide and 10% dextran sulphate. Probes were then

loaded on the denatured chromosomal slides. After overnight hybridisation, slides were washed and detected. FISH signals and the DAPI-banding pattern were recorded separately by taking photographs, and the assignment of the FISH mapping data with chromosomal bands was achieved by superimposing FISH signals with DAPI banded chromosomes.

Tissue Distribution of NAALAD-ases.

- 10 Northern Blot Analysis for NAALAD-ase III and IV.
Human MTN Northern blots containing 2 µg of poly(A)+ RNA derived from non-neuronal tissues were hybridised in ExpressHybTM hybridisation solution for 2 to 3 h at 68°C according to the manufacturer's instructions. A
- 15 546 bp NAALAD-ase III fragment was isolated from NAALAD-ase III/pcDNA3 (cl. 3.0) following digestion with EcoRI and BglII and a NAALAD-ase IV 526 bp fragment was isolated from NAALAD-ase IV/PCR2.1-TOPO (cl. 28.0) following digestion with PstI.
- 20 Radiolabelled probes were generated with these DNA fragments using a Rapid Multiprime Labelling kit and [³²P]-dCTP as label. Following removal of unincorporated label using a Micospin S-200 HR column (Pharmacia), 50 µl of hot denatured probe in 6 ml
- 25 ExpressHybTM hybridisation solution was incubated in a rotating shaker at 68 C overnight. Washes were 4 x 15 min RT in 2X SSC/0.05% SDS, 1 x 20 min wash at 50 C in 0.1x SSC/0.1% SDS and 2 x 20 min at 55 C in 0.1x SSC/0.1% SDS and 2 x 20 min. Blots were wrapped in
- 30 cling film, without drying and exposed to X-omat AR Scientific Imaging Film (Kodak Scientific Imaging Systems, Cambridge, UK) for 2 to 4 days at -70°C with two intensifying screens.
- 35 Gene Expression of NAALAD-ase I, II, III and IV by RT-

PCR analysis.

Oligonucleotide primers were designed for the specific amplification of a PCR fragment for each NAALAD-ase; NAALAD-ase I primers were NAALD1S3 5'- GGG AAA CAA ACA
5 AAT TCA GCG GC -3' and NAALD1AS3 5'- GTC AAA GTC CTG
GAG TCT CTC ACT GAA C -3' yielding a 341 bp product, NAALAD-ase II primers were NAALD2S7 5'- GAC CGG AGC
AAG ACT TCA GCC AG -3' and NAALD2AS7 5'- GTG TTG ATA
TGC GTT GGC CCA AG -3' yielding a 330 bp product,
10 NAALAD-ase III primers were NAALD3S4 5'- CAC TAA GAA
TAA GAA AAC AGA TAA GTA CAG C -3' and NAALD3AS4 5'-GAT
CAA CTT GTA TAA GTC GTT TAT GAA AAT CTG -3' yielding a
353 bp product and NAALAD-ase IV primers were NAALD4S1
5'- GCA GAA GAA CAA GGT GGA GTT GGT G -3' and
15 NAALD4AS1 5'- GCT TTG GAT CCA TGA CAG TCA TGG -3'
yielding a 336 bp product. Each primer set for each
NAALAD-ase was tested for its ability to specifically
amplify that NAALAD-ase and not to cross react in
amplification reactions with the other three forms.
20 PCR amplifications for human GAPDH were performed on
the same cDNA samples as positive controls using GAPDH
specific primers 5'-TGA AGG TCG GAG TCA ACG GAT TTG
GT-3' (sense primer) and 5'-CAT GTG GGC CAT GAG GTC
CAC CAC-3' (anti-sense primer), yielding a 1000 bp
25 fragment. These primers sets were used for PCR
amplifications on human multiple tissue cDNA (MTCTM)
panels normalised to the mRNA expression levels of six
different housekeeping genes. Human cDNAs from 15
brain regions were also prepared from mRNA and
30 normalised to the mRNA expression levels of three
different housekeeping genes, GAPDH, clathrin and
actin. Brain area mRNA was prepared starting from
carefully dissected tissue samples, using the
FastTrackR 2.0 kit (Invitrogen BV, Netherlands)
35 according to the manufacturer's instructions. 1 µg of

poly(A)+ RNA was reverse transcribed using oligo(dT)15 as a primer and 50 U of ExpandTM Reverse Transcriptase (Boehringer Mannheim, Germany) according to the manufacturer's instructions. Finally, normalised cDNA was prepared in a similar manner from 3 transformed prostate tumour cell lines and a prostate tumour. PCR reactions with NAALAD-ase or GAPDH specific primers were performed on 2 or 5 μ l of cDNA. PCR reactions were performed in a total volume of 50 μ l in 1X Advantage PCR buffer, 0.2 mM dNTP and 0.5 μ M of each PCR primer and 1 μ l Advantage Taq polymerase mix (95°C - 30 sec, 68°C - 1 min 30 s for 35 cycles). Upon completion of 25 cycles the PCR machine was paused at 80C, reaction tubes were removed and 15 μ l were removed from each PCR tube. Tubes were then returned to the machine and the cycling method continued. Aliquots were removed in a similar manner after 30 and 35 cycles. Each sample taken was analysed by 1.0 % agarose gel electrophoresis as previously described and images of the ethidium bromide stained gels were obtained using an Eagle Eye II Video system (Stratagene, La Jolla, CA, USA).

Results,

Molecular cloning and sequence analysis of NAALAD-ase II, III and IV.

NAALAD-ase II; Similarity searching of the Incyte LifeSeqTM database with the human, rat and mouse NAALAD-ase I sequences and with rat and partial human NAALAD-ase II protein sequences yielded 13 EST sequences, some of which were overlapping, encoding for at putatively novel protein sequences similar to NAALAD-ase I and II. DNA obtained from six of the most 5' clones were sequenced.

Incyte clone 4190746, from a cerebellar cDNA library,

contained sequences corresponding to NAALAD-ase II. However, since this clone also contained two segments of intronic sequence it was not suitable for further cloning experiments. PCR reactions were performed to amplify a PCR product containing the 3' half of the NAALAD-ase II coding region from a small intestine cDNA. To identify the remaining as yet unknown human 5' NAALAD-ase II sequence, 5' RACE PCR was performed on a number of cDNAs. Sequencing of the amplification products obtained from reactions using small intestine cDNA yielded a further 1344 bp fragment covering the complete coding sequence of NAALAD-ase II. The full cDNA sequence contained an open reading frame of 2223 bp encoding a protein of 740 amino acid residues with a calculated molecular mass of 80.6 kDa and an isoelectric point of 5.26 (Fig.. 1). The putative ATG translation start codon is in a favourable context for translation initiation (Kozak, 1989) with no ATG codons detected upstream. Analysis of the human NAALAD-ase II ORF predicted a type II integral membrane protein containing a single hydrophobic membrane-spanning domain extending from amino acid residues 6-27, with lysine residues bordering either side of the potential membrane-spanning domain (determined by the method of Kyte-Doolittle, 1982). There are 7 potential N-glycosylation sites (N X S/T), indicated in Fig. 1. The predicted protein sequences of human NAALAD-ase II was compared to that of rat using the alignment program Genedoc. Human NAALAD-ase II sequence was 78% identical and 87% similar to rat NAALAD-ase II. The two protein sequences were aligned using the ClustalW alignment program (EMBL, Heidelberg, Germany; Fig. 2).

NAALAD-ase II;; Incyte clones 1547649, 3448872,

3608639 and 1333965 contained sequences originating from the same gene that were similar to but not identical to NAALAD-ase I or II. Clone 3608639, from a lung carcinoma cDNA library, contained a 3110 bp DNA sequence, with a 2223 bp ORF coding for a 740 amino acid residue protein, which we termed NAALAD-ase III (Fig.. 4). Analysis of this open reading frame predicted a calculated molecular mass of 83.6 kDa and an isoelectric point of 8.53. The putative ATG translation start codon is in a favourable context for translation initiation (Kozak, 1989) and no ATG codons were detected upstream. NAALAD-ase III was predicted to be a type II integral membrane protein containing a hydrophobic membrane spanning domain extending from amino acid residues 8-31 (determined by the method of Kyte and Doolittle, 1982). There are also 7 potential N-glycosylation sites (N X S/T) indicated in Fig. 4.

NAALAD-ase IV; Sequencing of Incyte clone 2615389, derived from a gall bladder cDNA library, identified sequences originating from a gene with some similarity to NAALAD-ase I and II. Homology searching in the Incyte EST database using this new sequence information identified a contig of over 150 overlapping EST sequences spanning 1884 bp and containing an ORF of 1419 bp, plus 5' and 3' UTRs. Translation of the ORF predicted a 472 amino acid protein with calculated molecular mass of 51.9 kDa and an isoelectric point of 5.99 (Fig.. 5). We have named this protein NAALAD-ase IV. The putative ATG translation start codon is in a favourable context for translation initiation with no ATG codons detected upstream. Analysis of the NAALAD-ase IV sequence predicted a type II integral membrane protein containing a hydrophobic membrane spanning domain

extending from amino acid 3 to approximately 24 with lysine residues found on either side of the potential membrane-spanning domain (determined by the method of Kyte and Doolittle, 1982). There are 5 potential N-glycosylation sites (N X S/T) indicated in Fig. 5.

The predicted protein sequences of NAALAD-ase I, II, III and IV were compared to each other using the alignment program BESTFIT (Wisconsin package, Genetics Computer Group Software, Madison, Wisconsin, USA). The % identity and % similarity between each pair of sequences calculated by the Genedoc program is indicated in Table 1. NAALAD-ase I sequence was 67% identical to NAALAD-ase III, 35 % identical to NAALAD-ase II and 10% identical to NAALAD-ase IV. NAALAD-ase II and III were 37% identical. The four protein sequences for NAALAD-ase I, II, III and IV were aligned using the ClustalW alignment program (EMBL, Heidelberg, Germany) and are shown in Fig. 6. The highest regions of conservation between these proteins occur in the two predicted catalytic domains. A phylogram of NAALAD-ase I, II, III and IV was constructed using the GCG 'Distances' program with standard parameters and the 'Growtree' program with the UPGMA method and is depicted in Fig. 7. From this phylogram it is clear that NAALAD-ase I and III are the most closely related proteins with NAALAD-ase IV being a distant relative.

Two putative catalytic domains have been identified in NAALAD-ase I and II sequences by comparison to other peptidases (Bzegda et al., 1997; Shneider et al., 1997). Using multiple sequence alignments of NAALAD-ase I, II, III, and IV we have identified similar putative catalytic domains in human NAALAD-ase II, III and IV (see Fig. 6 and 8). The first catalytic

domain is related to bacterial and yeast Zn^{2+} dependent
peptidase domains (Bzegda et al., 1997). Alignments of
these peptidases with NAALAD-ases I, II, III and IV
are shown in Fig. 8. All five residues important in
5 Zn^{2+} binding are conserved in the NAALAD-ase family.
Additionally, a Glu residue thought to play a role in
catalysis is also present. The NAALAD-ase IV sequence,
although more distantly related, also shows
statistically significant similarity to the bacterial
10 and yeast amino-peptidase Zn^{2+} binding domain as
determined by the Blastp algorithm ($\sim 60\%$ similar,
 $p=10^{-5}$). The second catalytic domain is related to
members of the α/β hydrolase fold family of proteins
(Shneider et al., 1997). Interestingly, there is a
15 conserved nucleophile-acid-base alignment amongst
NAALAD-ase I, II and III that may be important in the
enzymes catalytic activity (Goossens et al., 1995).
Rat NAALAD-ase II has been suggested to belong to an
 α/β hydrolase fold family of peptidases, similar to
20 dipeptidyl peptidase IV (DPPIV). In contrast, NAALAD-
ase IV which is about 270 residues shorter than the
other NAALAD-ase proteins lacks this domain.

Alternative splicing of NAALAD-ase II.

25 In the course of the cloning and RT-PCR gene
expression analysis of NAALAD-ase II a number of
amplified PCR products were observed, isolated and
sequenced in order to identify possible splice
variants. We found both the splicing out of putative
30 exon sequences, as well as the presence of intronic
sequences, as judged by the presence of GT/AG
donor/acceptor sites, that were repeatedly amplified
from our cDNA preparations (Fig. 3). When performing
5' RACE amplifications deletions between bases 497-619
35 and 903-1007 were identified in small intestine and

colon that resulted in two in-frame amino acid deletions. In addition a 153 bp intron was found inserted at base 1094 resulting in an in-frame amino acid insertion of 51 amino acid residues. This
5 insertion is most likely an intron as it has the consensus GT/AG donor acceptor sites at its 5' and 3' ends respectively. In the 3' RACE amplification reactions several variants were also identified in amplifications from small intestine, colon, brain and
10 foetal brain. These consisted of either a deletion of bases 1525-1615 or a larger deletion between bases 1525-1697. Both these deletions resulted in frame shifts and the premature termination of the protein sequence. Finally, in every cDNA sample examined, two
15 intronic sequences were found to be inserted at either base 1697 and/or at base 1870. Inclusion of one or both of these intronic sequences into the ORF of NAALAD-ase II accounted for the unexpected amplified PCR products migrating at 420 and 500 bp in the RT-PCR
20 experiments. Introduction of one or both these sequences results in a frame shift of the amino acid coding sequence (Fig. 3).

Expression and functional activity of NAALAD-ases.
25 To determine if the newly identified NAALAD-ases had peptidase activity, mammalian expression constructs were transiently transfected into COS cells. Membranes from mock transfections were always prepared in parallel as negative controls. Expression of NAALAD-
30 ase I in COS cells was performed as a positive control to establish the working conditions of the assay. Fig. 9A shows the amount of [3H]-glutamate in dpms eluted off mini-columns following incubation of [3H]-NAAG with the recombinant NAALAD-ase I. Increasing [3H]-
35 glutamate is eluted from the column in a time

dependent manner. Activity was observed for both NAALAD-ase II (data not shown) and NAALAD-ase III (Fig. 9B) suggesting that these enzymes have a similar dipeptidase activity as NAALAD-ase I in this assay. No
5 NAALAD-ase IV dipeptidase activity was observed under these assay conditions. Addition of 30 μ M quisqualic acid to the reaction inhibited this activity by over 50% after 60 min (Fig. 9A and 9B). Inhibition curves with increasing concentrations of QA were also
10 performed (Fig. 9C) with IC₅₀ values of 1.2 x 10⁻⁵ M and 1.7 x 10⁻⁵ M for NAALAD-ase I NAALAD-ase III respectively.

A scintillation proximity assay (SPA) was also developed in order to facilitate screening of NAALAD-ase type activity using either the recombinant
15 proteins or membrane preparations from sources known to contain NAALAD-ase activity. A time course of [3H]-NAAG hydrolysis was initially performed using NAALAD-ase I rich LNCaP membrane preparations (Fig.. 10A).
20 Stopping the reaction with glycine buffer resulted in binding of the remaining [3H]-NAAG substrate to the SPA beads whereas hydrolysed [3H]-glutamate remained unbound. Therefore as the reaction proceeded and the substrate is hydrolysed, a decrease in signal bound to
25 the beads is observed. Under these SPA conditions, inhibition curves were performed on LNCaP membranes or recombinant NAALAD-ase I membrane preparations using increasing concentrations of QA (Fig.. 10B). The IC₅₀ values were 9.1 x 10⁻⁷ M for LNCaP preparations and 1
30 x 10⁻⁶ M for NAALAD-ase I preparations.

Chromosomal localisation.

The complete coding sequence of NAALAD-ase I was used as a probe for FISH analysis. Under the conditions
35 used, the hybridisation efficiency was approximately

71 % for this probe (among 100 checked mitotic
Fig.ures, 71 of them showed signals on one pair of the
chromosomes). DAPI banding was used to identify the
specific chromosome and an assignment between the
5 signal from the probe and the short arm of chromosome
11 was obtained. The detailed position was further
determined based upon summary data from 10 photographs
(Fig. 11). A weak hybridisation signal was also
detected in the region of 11q14.3 with low frequency.
10 With the mapping data obtained from NAALAD-ase III it
was concluded that this weak signal was a result of
cross-hybridisation (see later) and that NAALAD-ase I
is located solely at human chromosome 11, region
p11.21. Example of the mapping results are presented
15 in Fig. 11. For NAALAD-ase II, a 1059 bp fragment
(from pos. 1204-2262bp) was used as a probe for FISH
analysis. The hybridisation efficiency was
approximately 71 % and DAPI banding was used to
identify the signal to the long arm of chromosome 11.
20 The detailed position determined from 10 photographs
was on human chromosome 11, region q12 (Fig. 11). For
NAALAD-ase III, a 2.5 kb fragment (from pos. 1-2552
bp) was used as a probe for FISH analysis. The
hybridisation efficiency was approximately 74 % for
25 this probe. DAPI banding was used to identify the
signal to the long arm of chromosome 11. The detailed
position determined from 10 photographs was on human
chromosome 11, region q14.3-q21 (Fig. 11). For NAALAD-
ase IV, a 1539 bp fragment (from pos. 1-1539bp) was
30 used as a probe for FISH analysis. The hybridisation
efficiency was approximately 72 % for this probe. DAPI
banding was used to identify the signal to the long
arm of chromosome 8. The detailed position (Fig. 11)
determined from 10 photographs was on human chromosome
35 8, region q21.3.

Tissue distribution of NAALAD-ase III and IV as determined by Northern blot.

5 The tissue distribution of TrnR3 in different brain regions and in non-neuronal tissues was investigated using RT-PCR, Northern blot analysis and *in situ* hybridization. Northern blot analysis was performed on mRNA derived from different human tissues (Fig.. 12A). A NAALAD-ase III specific probe indicated the presence of transcripts in testis>>ovary, spleen>prostate, heart and placenta with no signal observed in other tissues. In testis, four transcripts were represented. 10 The most predominant transcript was of approximately 3.4 kb, consistent with the approximate expected size of a NAALAD-ase III message. Two transcripts of 2.4 and 4.4 kb respectively and a weaker transcript of about 7.5 kb were also present. In the other tissues the 3.4 kb transcript was the only signal detected, apart from ovary where a weak 7.5 kb signal could also be seen. The precise nature of these transcripts 15 awaits further elucidation but may be due to alternative splicing of the message. The nature of these messages is unclear.

20 A NAALAD-ase IV specific probe showed the ubiquitous presence of transcripts in all tissues examined with slightly less signal observed in brain, thymus and testis (Fig. 12B). The most predominant transcript was of approximately 2.2 kb, consistent with the approximate expected size of a NAALAD-ase IV message, with a faint band detected in some tissues at 4.4 kb.

30 Analysis of NAALAD-ase gene expression by RT-PCR. To further examine the detailed tissue distribution of all NAALAD-ases, PCR was performed on normalised cDNAs from 16 different tissues. Figure 13A shows the results from PCR reactions performed with NAALAD-ase I 35

specific primers, yielding amplification products of the expected size (~341 bp). Highest expression of NAALAD-ase I appeared to be in prostate. Rank order of expression after 25 cycles was prostate >> liver and kidney > small intestine > brain, spleen, with no product amplification observed in the other tissues. At 30 cycles amplification products could be seen in most other tissues with the exception of muscle, blood and thymus in which products could only be observed after 35 cycles of amplification. NAALAD-ase II specific primers yielded a 330 bp amplification product of the expected size, as well as two products migrating with slightly higher sizes of 420 bp and 500 bp (Fig. 13B). NAALAD-ase II expression was highest in small intestine, spleen and testis with PCR products detected after 25 cycles of amplification, whilst products in heart, ovary, colon, blood and prostate could be seen after 30 cycles. All tissues exhibited some amplification products following 35 cycles with brain and muscle showing the lowest levels. The 420 and 500 bp bands were due to amplification of NAALAD-ase II sequences containing one or two intronic sequences that were commonly found in all our amplification reactions (see earlier). NAALAD-ase III specific primers yielded a 353 bp amplification product of the expected size (Fig. 13C). NAALAD-ase III expression was highest in ovary, testis and spleen with PCR products detected after 25 cycles of amplification. After 30 cycles amplification products could be detected from all tissue cDNAs apart from lung, muscle, blood and thymus in which a product could only be seen after 35 cycles. These results are in good accordance with the expression data obtained with the MTN Northern blots. NAALAD-ase IV specific primers yielded a 336 bp amplification product of the

expected size (Fig. 13D). NAALAD-ase IV expression was low in all tissues with amplification products only detected after 35 cycles. Expression levels were comparable in all the cDNAs apart from brain, lung, muscle and thymus in which no amplification products were clearly visible. Control amplification reactions using GAPDH specific primers demonstrated comparable levels of amplification products for each cDNA (Fig. 13E). Comparison of the relative abundance between the four messages was also possible from these experiments, since the same cDNAs were used for each set of amplifications. Abundance of NAALAD-ase I message was greater than NAALAD-ase III which was greater than NAALAD-ase II, as judged by the relative amount of amplification products detected at 25 and 30 cycles. NAALAD-ase IV was expressed at the lowest levels, with PCR products detected only after 35 cycles of amplification.

PCR reactions were also performed with the same NAALAD-ase primers as in the above experiments, on 13 different brain cDNAs normalised to the expression levels of three different housekeeping genes. NAALAD-ase I specific amplification products were detected with highest levels in ventral striatum and brain stem after 30 cycles. After 35 cycles NAALAD-ase I specific amplification products could clearly be detected in all brain areas studied (Fig. 14A). NAALAD-ase II specific primers yielded a 330 bp amplification product of the expected size, as well as a product migrating at a higher size of 500 bp (Fig. 14B). Amplification of the 500 bp product was observed after 35 cycles in brain stem, amygdala, thalamus, ventral striatum and to a lesser extent in striatum and hippocampus, whereas the 330 bp product of the expected size was only seen in brain stem and ventral

striatum. NAALAD-ase III specific primers yielded a 353 bp amplification product of the expected size (Fig. 14C). Amplification products were observed after 30 cycles in striatum, parietal cortex and ventral striatum with lower levels of amplification product detected in hippocampus, brain stem, putamen and superior colliculus. After 35 cycles the presence of NAALAD-ase III specific products could be detected in all cDNAs apart from inferior colliculus. NAALAD-ase IV specific products could be detected in all brain cDNAs studied after 30 cycles with little difference observed in expression pattern. Levels of amplification products were slightly higher in striatum, hippocampus, brain stem and ventral striatum (Fig. 14D). Control amplification reactions using GAPDH specific primers demonstrated comparable levels of amplification products for each cDNA apart from brain stem which yielded relatively more GAPDH specific product (Fig. 14E). Overall expression of NAALAD-ase II appears to be lower in these brain areas relative to NAALAD-ase I, III and IV. Finally, NAALAD-ase expression was investigated in cDNAs prepared from either prostate tumour cell lines or a prostate tumour, that had been normalised against three different housekeeping genes. NAALAD-ase I expression was highest in LNCaP and prostate tumour (Fig. 15A). Amplification products were also detected in PC-3 cDNA after 30 and 35 cycles but not in DU145 cDNA. The 330 bp NAALAD-ase II product was detected in highest amounts in cDNA from prostate tumour and less in PC-3 and DU145 samples after 35 cycles (Fig. 15B). Interestingly, in all samples apart from prostate tumour, the higher 500 bp amplification product could be detected. NAALAD-ase III expression was higher in LNCaP than prostate tumour. A faint amplification

product could also be detected in PC-3 after 35 cycles but not in DU145 cDNA (Fig. 15C). NAALAD-ase IV expression was highest in LNCaP and prostate tumour (Fig. 15D). Representative amplifications with GAPDH
5 primers are also shown (Fig. 15E).

Discussion

NAALAD-ase I is a type II membrane glycoprotein, with carboxypeptidase activity and sequence similarity to
10 the transferrin receptor (54% identical, Israeli et al., 1993). It may be important in the progression of prostate cancer since it is highly expressed in prostate tumours, and in the CNS it may have an important role in modulating neuronal glutamatergic activity. In this report we describe the
15 identification and cloning of human cDNAs encoding distinct NAALAD-ase enzymes, that represent an expansion of this family of proteins. We have identified In this report two novel members of the NAALAD-ase enzyme family, NAALAD-ase III and NAALAD-
20 ase IV and determined the full coding sequence of human NAALAD-ase II. The degree of divergence of NAALAD-ase II, III and IV at the amino acid sequence level indicates that these are not variants of existing known genes but new and distinct members of
25 this family. As has been determined for NAALAD-ase I, aFeatures common to all three receptors include a hydrophobic leader sequence at the amino-terminus suitable for secretion, and the presence of several possible N-glycosylation sites. 11 three novel NAALAD-
30 ases contain a single hydrophobic region of approximately 21 to 24 amino acid residues at their amino-terminus, which is likely membrane spanning, a short intracellular domain and a large globular extracellular domain. This structure is typical of
35 type II integral membrane proteins and is common

amongst membrane bound hydrolases (Kenny et al., 1982). Sequence alignment comparisons of the four proteins suggests that NAALAD-ase III is most closely related to NAALAD-ase I (81% similar at the amino acid level), whereas NAALAD-ase II is equally similar to both NAALAD-ase I and III (54% similar to both). NAALAD-ase IV is distantly related to the other members of this family.

Analysis of NAALAD-ase II sequences using different PCR primers sets revealed the presence of multiple splice variants and isoforms. The biological significance of alternatively spliced isoforms of NAALAD-ase II awaits further clarification. However, the spliced out residues may affect levels of glycosylation and more importantly the conformation and activity of the protein. Furthermore, inclusion of three different intronic sequences identified in amplification reactions from numerous cDNAs, result in either an in-frame addition of a proline rich, 51 amino acid residue, sequence close to the putative catalytic zinc binding domain, or in frame shifts and subsequent premature termination of the NAALAD-ase II protein. These premature terminations result in elimination of the predicted nucleophile-acid-base arrangement of the putative catalytic site found in α/β hydrolases and will result in a gross change in the proteins conformation (see Fig. 3). Inclusion of two of the intronic DNA sequences, resulting in frame shifts and premature protein termination, were identified in nearly all the cDNAs studied, suggesting that expression of these sequences may be used to regulate the levels of active protein. Expression studies using these different splice variants will help to understand the structure/activity relationship of NAALAD-ase II and help identify which amino acid

residues within the protein are important for enzymatic activity. A soluble splice variant of NAALAD-ase I has been reported, lacking the first 40 amino acid residues, including the membrane spanning domain (Su et al., 1995). An increased ratio of full length protein to truncated protein was observed in malignant prostate tissues, suggesting that expression of alternative splice variants may correlate with tumour progression. Further work will determine if NAALAD-ase II is involved in oncogenesis and if these splice variants effect tumour progression to differing degrees.

In human tissues, NAALAD-ase I mRNA was highly expressed in prostate but also in liver, kidney and small intestine and to a lesser extent in brain and spleen. This in partial agreement with Northern blot results obtained in rat in which the highest levels of expression were found to be in brain and kidney but not liver or spleen (prostate was not studied; Bzedga et al. 1997; Carter et al., 1998). NAALAD-ase II expression by RT-PCR was highest in spleen, small intestine and testis. Schneider and colleagues (1997) identified a rat ileal 100 kD protein (I-100; rat NAALAD-ase II) expressed in rat ileum and human small intestine as seen by Northern blot, but not in testis or spleen. Discrepancies between NAALAD-ase I and II expression patterns could be attributed either to species differences or to the increased sensitivity of mRNA detection by RT-PCR compared to Northern blot hybridisation. NAALAD-ase III mRNA expression was studied both by Northern blot and RT-PCR experiments. Expression experiments using these methods yielded comparable results with the highest levels of mRNA expression being observed in testis and ovary, with less message in spleen, placenta and heart. A number

of differently sized transcripts were observed that hybridised to the the NAALAD-ase III probe, suggesting the presence of multiple isoforms. It remains to be determined whether these are due to physiologically relevant smaller transcripts or due to problems with the integrity of the mRNA. NAALAD-ase IV mRNA expression was also studied using both techniques and expression data were consistent between the experiments. NAALAD-ase IV expression appeared to be more evenly distributed amongst tissues, with the highest expression in ovary and little or no expression in brain, lung, thymus and testis. Analysis of NAALAD-ase expression in different human brain areas by RT-PCR suggests that NAALAD-ase II is expressed at lowest levels in these brain areas. The amplification products (except those from ventral striatum) migrating at 500 bp were of a higher than expected size and found to contain intronic DNA sequences (see earlier). NAALAD-ase I, III and IV were expressed at the mRNA level in almost all brain areas studied to differing degrees, with highest expression observed in brain stem, striatum, ventral striatum and putamen. In addition, NAALAD-ase III expression appeared relatively more abundant in frontal and parietal cortex. RT-PCR experiments performed on various prostate tumour cell lines and on a prostate tumour sample agreed with reports showing that NAALAD-ase I is expressed in highest amounts in the androgen receptor positive LNCaP cell line (Israeli et al., 1993). Some amplification products were also observed in PC-3 cells after 30 and 35 cycles of amplification, in contrast to previous studies detecting no message in the androgen receptor negative cells DU-145 and PC-3 by Northern blotting (Israeli et al. 1993). Since low levels of NAALAD-ase activity have been observed

in these latter two cell lines, some expression of these enzymes, including NAALAD-ase I, may occur (Pinto et al., 1996). NAALAD-ase III was also expressed in LNCaP cells at the highest levels, whereas NAALAD-ase IV expression was highest in PC-3 cells. Interestingly, NAALAD-ase II amplification products from LNCaP cells were of a higher than expected size and contained intronic sequences, in contrast to the product observed from prostate tumour. PC-3 and DU-145 cells yielded both the high and low sized amplification products.

Carboxypeptidase activity of NAALAD-ase I has been demonstrated against two classes of substrate *in vitro*. The first is based on the enzymes ability to hydrolyse α -linked acidic peptides such as N-acetyl-aspartyl glutamate (NAAG) or α -glutamylglutamate and the second is based on its ability to hydrolyse γ -linked peptides such as folyl-poly- γ -glutamate or γ -glutamylglutamate. In this study we have expressed NAALAD-ase II, III and IV in COS cells and studied the ability of these recombinant preparations to hydrolyse NAAG. NAALAD-ase I, II and III all showed activity against this substrate that was inhibited by QA. QA IC50s for NAALAD-ase I and III were determined to be about 10 μ M for both proteins, comparable to that described for NAALAD-ase activity in lysed synaptosomal membranes (Robinson et al., 1987). The α -linked NAALAD-ase hydrolysis of NAAG was initially characterised in membrane preparations from brain and shown to be QA sensitive. Activity was increased by divalent cations and inhibited by divalent metal chelators or general metalloprotease inhibitors (Robinson et al., 1987; Blakely et al, 1988). It was subsequently shown that the enzymatic activity associated with prostate specific marker

(PSM, Carter et al., 1996) and rat brain
carboxypeptidase II (Luthi-Carter et al., 1998)
following transfection experiments, exhibited similar
enzyme characteristics to those observed for the QA
5 sensitive NAALAD-ase activity. It has previously been
shown that alignment of NAALAD-ase I with zinc amino
peptidases from yeast and bacteria identifies a
domain of conserved sequences involved in the co-
ordination of two zinc ions located in the catalytic
10 site (Bzedga et al., 1997, Rawlings et al., 1997). The
basis for these observations comes from the three
dimensional crystal structure of the *Aeromonas*
proteolytica zinc aminopeptidase in which residues
His 379, Asp 389, Glu 427, Asp 455, His 555 are
15 thought to be involved in the binding of two zinc ions
and Glu426 is proposed to be a base residue important
in catalysis (Chevrier et al., 1994, 1996). Alignment
of the newly identified NAALAD-ases with these amino
peptidases (Fig. 8) shows that these zinc binding
20 domains are conserved, suggesting that they may all
share the same catalytic mechanism. Bzedga and
colleagues (1997) have further suggested that NAALAD-
ase I may be a member of the α/β hydrolase fold family
because of its dipeptidyl IV and acylaminoacyl
25 peptidase activity, although sequence homology
alignments have shown these proteins to be clearly
distinct. However, a hypothesised catalytic site
arrangement of nucleophile, acid and base (Ser 623,
Asp 663 and His 686) is conserved and found in NAALAD-
30 ase I, II and III but absent from NAALAD-ase IV (see
Fig. 6). The second γ -linked enzymatic activity of
NAALAD-ase I to peptides such as pteroyl- γ -glutamate
(folate hydrolase activity) has been found at high
levels in the brush border of human small intestine
35 (Chandler et al., 1986; Pinto et al., 1996). In

addition, carcinoma cells transfected with NAALAD-ase I show increased folate hydrolase activity and the ability to progressively liberate glutamate from methotrexate triglutamate by hydrolysis of γ -glutamyl linkages (Pinto et al., 1996). Indeed, a correlation has been observed between increased pteroyl hydrolase activity and methotrexate resistance in tumours (Banerjee et al., 1986?), suggesting that modulating NAALAD-ase activity may be important in developing improved or novel cancer treatments. It will be interesting to determine if these novel NAALAD-ase enzymes with their common secondary structure also have a dual peptidase activity as has been described for NAALAD-ase I, and how their structural relationship reflects their activity.

Using FISH analysis to determine the chromosomal localisation of NAALAD-ase I, II and III we observed a symmetrical fluorescent signal on chromosome 11, at p11.21 for NAALAD-ase I, q12 for NAALAD-ase II and between q14.3-q21 for NAALAD-ase III. In our FISH studies a NAALAD-ase I probe revealed two hybridisation signals, one hybridising at 11p11.21 and another weakly at 11q14.3. This is similar to the results obtained by Leek and colleagues (1995) who observed two hybridisation signals at 11p11.2 and 11q13.5. Having now localised NAALAD-ase III to 11q14.3-q21, it is clear that the second signal is due to cross hybridisation of the NAALAD-ase I probe with the NAALAD-ase III locus. The chromosomal localisation of NAALAD-ase IV revealed a symmetrical fluorescent signal on chromosome human chromosome 8 at q21.3, suggesting that this gene is distantly related to the other three genes located on chromosome 11. Chromosome 11 contains a number of genetic disease loci in these regions, including vitreoretinopathy (11q13-q23),

xeroderma pigmentosa (11q12-q13), atopy (11q12-q13) and perhaps more interestingly a tumour suppression locus (11p11.2-p11.13) involved in rat prostate carcinoma. Introduction of this portion of the chromosome into highly metastatic rat prostatic cells was able to suppress cancer metastases without suppression of the in vivo growth rate or tumourgenicity of the cells (Ichikawa et al., 1992). Since it has been shown that NAALAD-ase I expression increases with decreasing androgen levels, it is possible that current prostate cancer treatments involving androgen level reduction (e.g. orchidectomy) may work at least in part through alteration of NAALAD-ase expression (Israeli et al., 1994). The enzymatic activity of NAALAD-ase I is similar to that of NAALAD-ase II and III, so it is also conceivable by analogy that these enzymes may also have a role to play in tumour suppression. Interestingly, region 11q13-q23 has also been identified as a region with tumour suppressor activity using tumourgenic HeLa/fibroblast hybrids (Misra and Srivatsan, 1989). In addition, in a systematic analysis of primary cervical carcinomas, region 11q22-q24, was shown to contain tumour suppressor activity (Hampton et al., 1994). These latter two tumour suppressing regions on the long arm of chromosome 11, cover the gene loci of NAALAD-ase II and NAALAD-ase III identified in this study. However, it should be noted that mapping of these tumour suppressing activities to these three chromosomal regions, in no way establishes if any of the identified NAALAD-ases are capable of having tumour or metastatic suppressing activity. α -NAAG is one of the most abundant peptides in the brain, being present in millimolar concentrations in certain brain regions (see Coyle, 1997). Clarification

of the physiological role of NAAG has been difficult as it is co-localised with L-glutamate, but it does fulfil some of the criteria for a neurotransmitter. It is localised in synaptic vesicles, is released in Ca²⁺ dependent manner from nerve terminals and is rapidly inactivated by a QA sensitive enzyme activity first identified in brain membrane preparations (Robinson et al., 1986, 1987). Cloning experiments revealed the sequence identity of this enzymatic activity to be derived from a previously identified protein, PSM (or NAALAD-ase I; Carter et al., 1996). In this study, NAALAD-ases II and III but not NAALAD-ase IV, have been shown to be able to hydrolyse [3H]-NAAG to N-acetylaspartate and glutamate as has been previously shown with NAALAD-ase I. Given the localisation of these NAALAD-ases in prostate and ovary, as well as other peripheral tissues, it is quite possible that these enzymes may modulate local extracellular glutamate levels in these tissues. For example it is known that substantial amounts of glutamate are present in seminal fluid. In the CNS, NAAG has been shown to act as a partial agonist at the NMDA receptor but not at AMPA or kainate receptors (Vallivullah et al., 1994; Puttfarcken et al., 1993; Sekiguchi et al., 1992) and to attenuate NMDA or glutamate induced neurodegeneration (Bruno et al., 1998). Furthermore, it has also been shown that addition of NAAG to hippocampal slices or mixed cortical cultures results in neuroprotection following addition of excitotoxins via a mechanism distinct from its action on the NMDA receptor. In this case it is postulated that NAAG, which is poorly transported or actively taken up, diffuses from the synaptic cleft and binds as an agonist to type II metabotropic glutamate receptors (mGluR), such as mGluR3, leading to reduced

glutamatergic neurotransmission (Wroblewska et al., 1993 and 1997). This hypothesis has been further tested using the mGluR3 antagonist, ethylglutamate, which eliminated the neuroprotective actions of NAAG (Bruno et al., 1998). Although the intact peptide has predominantly inhibitory actions, another possible mechanism of excitotoxic induced cell death may arise from the aberrant catabolism of NAAG by NAALAD-ases. With endogenous levels of NAAG being so high in brain, catabolism of NAAG by NAALAD-ases should in theory be a rich source of glutamate, which if not properly regulated may result in excitotoxic effects. Indeed abnormal levels of NAAG or NAALAD-ase activity have been reported in epilepsy (Myerhoff et al., 1985 and 1992), schizophrenia (Tsai et al., 1995), ALS (Tsai et al., 1991), Alzheimer's disease (Passani et al., 1997a; Jaarsma et al., 1994) and stroke (Sager et al., 1995). Inhibition of NAALAD-ase activities with specific inhibitors may be useful in treating ischemic induced neurodegeneration or other neurodegenerative disorders involving abnormalities in glutamate neurotransmission, such as Alzheimer's disease, schizophrenia or ALS (Passani et al., 1997b). In vitro at least, the NAALAD-ase (carboxypeptidase) inhibitor, 2-(phosphonomethyl)pentanedioic acid, inhibited toxicity induced by the carboxypeptidase cleavage of folic acid hexaglutamate (Slusher et al., 1997).

References.

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and
Lipman, D.J. (1990). Basic local alignment search
tool. *J. Mol. Biol.*, 215, 403-410.
- 5 Bruno, V., Wroblewska, B., Wroblewski, J.T., Fiore,
L. and Nicoletti, F. (1998) Neuroprotective activity
of n-acetylaspartylglutamate in cultured cortical
cells. *Neurosci.*, 85(3), 751-757.
- 10 Burlina, A.P., Skaper, S.D., Mazza, M.R., Ferrari,
V., Leon, A. and Burlina, A.B. (1994) N-
acetylaspartylglutamate selectively inhibits neuronal
responses to N-methyl-D-aspartic acid in vitro. *J.*
15 *Neurochem.*, 63(3), 1174-7, 1994.
- Bzdega, T., Turi, T., Wroblewska, B., She, D.,
Chung, H.S., Kim, H. and Neale, J.H. (1997).
Molecular cloning of a peptidase against N-
20 acetylaspartylglutamate from a rat hippocampal cDNA
library. *J. Neurochem.*, 69, 2270-2277.
- Carter, R.E. Feldman, A.R. and Coyle, J.T. (1996)
Prostate-specific membrane antigen is a hydrolase with
25 substrate and pharmacologic characteristics of a
neuropeptidase. *P.Natl. Acad.Sci. USA.*, 93(2), 749-
753.
- Cassidy, M. and Neale, J.H. (1993) N-
30 acetylaspartylglutamate catabolism is achieved by an
enzyme on the cell surface of neurons and glia.
Neuropeptides., 24(5), 271-278.
- Chandler, C.J., Wang, T.T. and Halsted, C.H. (1996)
35 Pteroylpolyglutamate hydrolase from human jejunal

brush borders. Purification and characterization. J. Biol.Chem., 261(2), 928-933.

5 Chevrier, B., D'Orchymont, H., Schalk, C., Tarnus, C. and Moras, D. (1996) The structure of the *Aeromonas proteolytica* aminopeptidase complexed with a hydroxamate inhibitor. Involvement in catalysis of Glu151 and two zinc ions of the co-catalytic unit. Eur. J. Biochem., 237(2), 393-398.

10 Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.M., Moras, D. and Tarnus, C. (1994) Crystal structure of the *Aeromonas proteolytica* aminopeptidase, a prototypical member of the co-catalytic zinc enzyme family. Structure, 2, 283-291.

20 Coyle, J.T. (1997) The nagging question of the function of N-acetylaspartylglutamate. Neurobio. Disease., 4(3-4), 231-238.

25 Ghose, S., Wroblewska, B., Corsi, L., Grayson, D.R., De Blas, A.L., Vicini, S. and Neale, J.H. (1997) N-acetylaspartylglutamate stimulates metabotropic glutamate receptor 3 to regulate expression of the GABA(A) α_6 sub-unit in cerebellar granule cells. J. Neurochem., 69(6), 2326-2335.

30 Hampton, G.M., Penny, L.A., Baergen, R.N., Larson, A., Brewer, C., Liao, S., Busby-Earle, R.M., Williams, A.W., Steel, C.M., Bird C.C., Stanbridge, E.J. and Evans, G.A. (1994) Loss of heterozygosity in sporadic human breast carcinoma, a common region between 11q22 and 11q23.3. P.Natl. Acad.Sci. USA., 91, 6953-6957.

35

- Ichikawa, T., Ichikawa, Y., Dong, J., Hawkins,
A.L., Griffin, C.A., Isaacs, W.B., Oshimura, M.,
Barrett, J.C. Isaacs, J.T. (1992) Localization of
metastasis suppressor gene(s) for prostatic cancer to
5 the short arm of human chromosome 11. Cancer Res.,
52(12), 3486-3490.
- Israeli, R.S., Powell, C.T., Corr, J.G., Fair, W.R.
and Heston, W.D. (1994) Expression of the prostate-
10 specific membrane antigen. Cancer Res., 54(7), 1807-
1811.
- Israeli, R.S., Powell, C.T., Fair, W.R. Heston,
W.D. (1993) Molecular cloning of a complementary DNA
15 encoding a prostate-specific membrane antigen. Cancer
Res., 53(2), 227-230. J. Neurochem., 69(1), 174-181.
- Jaarsma, D., Veenma-van der Duin, L. and Korf, J.
(1994) N-acetylaspartate and N-acetylaspartylglutamate
20 levels in Alzheimer's disease post-mortem brain
tissue. J. Neurol. Sci., 127(2), 230-233.
- Kenny, A.J. and Maroux, S. (1982) Topology of
microvillar membrane hydrolases of kidney and
25 intestine. Physiol. Revs. 62(1), 91-128.
- Kozak, M. (1989) The scanning model for translation,
an update. J. Cell Biol., 108, 229-241.
- 30 Kyte, J. and Doolittle, R.F. (1982) A simple method
for displaying the hydropathic character of a protein.
J. Mol. Biol., 157(1), 105-132.
- Leek, J., Lench, N., Maraj, B., Bailey, A., Carr,
35 I.M., Andersen, S., Cross, J., Whelan, P.,

- MacLennan, K.A., Meredith, D.M. and Markham, A.F.
(1995) Prostate-specific membrane antigen, evidence
for the existence of a second related human gene.
Brit. J. Cancer, 72(3), 583-588.
- 5 Luthi-Carter, R., Berger, U.V., Barczak, A.K.,
Enna, M. Coyle, J.T. (1998) Isolation and expression
of a rat brain cDNA encoding glutamate
carboxypeptidase II. P.Natl. Acad.Sci. USA., 95(6),
10 3215-3220.
- Meyerhoff, J.L., Carter, R.E., Yourick, D.L.,
Slusher, B.S. and Coyle, J.T. (1992) Activity of a
NAAG-hydrolyzing enzyme in brain may affect seizure
15 susceptibility in genetically epilepsy-prone rats.
Epilepsy Research (Supp.), 9, 163-172.
- Meyerhoff, J.L., Robinson, M.B., Bixler, M.A.,
Richards, S.S. and Coyle, J.T. (1989) Seizures
20 decrease regional enzymatic hydrolysis of N-acetyl-
aspartylglutamate in rat brain. Brain Res., 505(1),
130-134.
- Meyerhoff, J.L., Koller, K.J., Walczak, D.D. and
25 Coyle, J.T. (1985) Regional brain levels of N-acetyl-
aspartyl-glutamate, the effect of kindled seizures.
Brain Res., 346(2), 392-396.
- Misra, B.C. and Srivatsan, E.S. (1989) Localization
30 of HeLa cell tumor-suppressor gene to the long arm of
chromosome 11. Am. J. Hum. Genet., 45(4), 565-577.
- Orlando, L.R., Luthi-Carter, R., Standaert, D.G.,
Coyle, J.T., Penney, J.B Jr. and Young, A.B. (1997)
35 N-acetylaspartylglutamate (NAAG) protects against rat

striatal quinolinic acid lesions in vivo. Neurosci. Letts., 236(2), 91-94.

- 5 Passani, L.A., Vonsattel, J.P. and Coyle, J.T. (1997) Distribution of N-acetylaspartyl-glutamate immunoreactivity in human brain and its alteration in neurodegenerative disease. Brain Res., 772(1-2), 9-22.
- 10 Passani, L.A., Vonsattel, J.P., Carter, R.E. and Coyle, J.T. (1997) N-acetylaspartyl-glutamate, N-acetylaspartate, and N-acetylated alpha-linked acidic dipeptidase in human brain and their alterations in Huntington and Alzheimer's diseases. Mol. Chem. Neuropathol., 31(2), 97-118.
- 20 Pinto, J.T., Suffoletto, B.P., Berzin, T.M., Qiao, C.H., Lin, S., Tong, P.T., May, F., Mukherjee, B. and Heston, W.D. (1996) Prostate-specific membrane antigen, A novel folate hydrolase in human prostatic carcinoma cells. Clin. Cancer Res., 2, 1445-1451.
- 25 Puttfarcken, P.S., Handen, J.S., Montgomery, D.T., Coyle, J.T. and Werling, L.L. (1993) N-acetyl-aspartylglutamate modulation of N-methyl-D-aspartate-stimulated [3H] norepinephrine release from rat hippocampal slices. J. Pharmacol. Exp. Ther., 266(2), 796-803.
- 30 Rawlings, N.D. and Barrett, A.J. (1997) Structure of membrane glutamate carboxypeptidase. Biochim. et Bioph. Acta., 1339(2), 247-252.
- 35 Robinson, M.B., Blakely, R.D., Couto, R. and Coyle, J.T. (1987) Hydrolysis of the brain dipeptide N-

acetyl-L-aspartyl-L-glutamate. Identification and characterization of a novel N-acetylated alpha-linked acidic dipeptidase activity from rat brain. J. Biol. Chem., 262(30), 14498-14506.

5

Robinson, M.B., Blakely, R.D. and Coyle, J.T. (1986) Quisqualate selectively inhibits a brain peptidase which cleaves N-acetyl-L-aspartyl-L-glutamate in vitro. Eur. J. Biochem., 130(3), 345-347.

10

Sager, T.N., Laursen, H. and Hansen, A.J. (1995) Changes in N-acetyl-aspartate content during focal and global brain ischemia of the rat. J. Cerebr. Blood F. Met., 15(4), 639-646.

15

Sekiguchi, M., Wada, K. and Wenthold, R.J. (1992) N-acetylaspartylglutamate acts as an agonist upon homomeric NMDA receptor (NMDAR1) expressed in Xenopus oocytes. FEBS Letts., 311(3), 285-289.

20

Serval, V., Barbeito, L., Pittaluga, A., Cheramy, A., Lavielle, S. and Glowinski, J. (1990) Competitive inhibition of N-acetylated-alpha-linked acidic dipeptidase activity by N-acetyl-L-aspartyl-beta-linked L-glutamate. J. Neurochem., 55(1), 39-46.

25

Shneider, B.L., Thevananther, S., Moyer, M.S., Walters, H.C., Rinaldo, P., Devarajan, P., Sun, A.Q., Dawson, P.A. and Ananthanarayanan M. (1997) Cloning and characterisation of a novel peptidase from rat and human ileum. J. Biol. Chem., 272(49), 31006-31015.

30

Slusher, S., Thomas A.T., Jackson, P.F., Tiffany, C.W. and Suzdak, P.D. (1997) Folate polyglutamate toxicity

35

in rat cortical neurones is attenuated by NAALAD-ase inhibition. Soc. Neurosci. Abst., 898.7.

- 5 Su, S.L., Huang, I.P., Fair, W.R., Powell, C.T. and Heston, W.D. (1995) Alternatively spliced variants of prostate-specific membrane antigen RNA, ratio of expression as a potential measurement of progression. Cancer Res., 55(7), 1441-1443.
- 10 Tsai, G., Passani, L.A., Slusher, B.S., Carter, R., Baer, L., Kleinman, J.E. and Coyle, J.T. (1995) Abnormal excitatory neurotransmitter metabolism in schizophrenic brains. Arch. Gen. Psych., 52(10), 829-836.
- 15 Tsai, G.C., Stauch-Slusher, B., Sim, L., Hedreen, J.C., Rothstein, J.D., Kuncl, R. and Coyle, J.T. (1991) Reductions in acidic amino acids and N-acetylaspartylglutamate in amyotrophic lateral
- 20 sclerosis CNS. Brain Res., 556(1), 151-156.
- Valivullah, H.M., Lancaster, J., Sweetnam, P.M. and Neale, J.H. (1994) Interactions between N-acetylaspartylglutamate and AMPA, kainate, and NMDA
- 25 binding sites. J. Neurochem., 63(5), 1714-1719.
- Wroblewska B., Wroblewski J.T., Saab O.H., Neale J.H. (1993) N-acetylaspartylglutamate inhibits forskolin-stimulated cyclic AMP levels via a metabotropic
- 30 glutamate receptor in cultured cerebellar granule cells. J. Neurochem., 63(3), 943-948.
- Wroblewska, B., Wroblewski, J.T., Pshenichkin, S., Surin, A., Sullivan, S.E. and Neale, J.H. (1997) N-
- 35 acetylaspartylglutamate selectively activates mGluR3

receptors in transfected cells. J.Neurochem., 69,
174-181.

Table 1:

	NAALAD-I	NAALAD-II	NAALAD-III	NAALAD-IV
NAALAD-I	100	54	81	21
NAALAD-II	35	100	54	17
NAALAD-III	67	37	100	21
NAALAD-IV	10	9	11	100

Comparison of the predicted protein sequences for NAALAD-ase I-IV. The sequences were compared two by two using the alignment program BESTFIT (GCG package). The obtained values for % identity (upper part of the table) and % similarity (lower part of the table) are shown.

List of used abbreviations.

BLAST	basic local alignment search tool
bp	base pairs
CMV	cytomegalovirus
DMEM	defined minimal essential medium
EST	expressed sequence tag
FISH	fluorescent in situ hybridisation
GAPDH	glyceraldehyde-3-phosphate
dehydrogenase	
I100	ileal 100 kDa protein
IPTG	isopropyl- β -D-thiogalactoside
MEM	minimal essential medium
NAALAD-ase	N-acetylated alpha-linked acidic
dipeptidase	
NAAG	N-acetyl-L-aspartyl-L-(3,4)-glutamate
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction

	PSM	prostate specific marker
	QA	quisqualate
	RACE	rapid amplification of cDNA ends
	RT-PCR	reverse transcription polymerase chain
5	reaction	
	SDM	site directed mutagenesis
	SPA	scintillation proximity assay
	SSC	sodium chloride sodium citrate
	TAE	Tris acetate EDTA
10	X-Gal	5-bromo-4-chloro-3-indolyl- β -D-
	galactopyranoside	

Claims

1. A cDNA molecule encoding a human NAALAD-ase II protein having the amino acid sequence illustrated in Figure 1 or a functional equivalent, derivative or bioprecursor thereof.
5
2. A cDNA molecule according to claim 1 which cDNA comprises the sequence of nucleotides illustrated in Figure 1.
10
3. A cDNA molecule encoding any of the splice variants of human NAALAD-ase II proteins having the amino acid sequences illustrated in Figure 3.
15
4. A cDNA molecule according to claim 3 which DNA molecule comprises the sequence of nucleotides illustrated in Figure 1 having respective nucleotide deletions or insertions encoding the respective splice variants of human NAALAD-ase II illustrated in Figure 3.
20
5. A cDNA molecule encoding a NAALAD-ase III protein having the amino acid sequence illustrated in Figure 4 or a functional equivalent, derivative or bioprecursor thereof.
25
6. A cDNA molecule according to claim 5 which cDNA molecule comprises the nucleotide sequence illustrated in Figure 4.
30
7. A cDNA molecule encoding a NAALAD-ase IV protein having the amino acid sequence illustrated in Figure 5 or a functional equivalent, derivative or bioprecursor thereof.
35

8. A cDNA molecule according to claim 7 having the nucleotide sequence illustrated in Figure 5.

5 9. A nucleic acid molecule capable of hybridising to the cDNA molecules according to any of claims 1 to 8 under high stringency conditions.

10 10. A human NAALAD-ase II protein having an amino acid sequence encoded by the nucleotide sequence illustrated in Figure 1 or a functional equivalent or derivative thereof.

15 11. A human NAALAD-ase II protein according to claim 10 having the amino acid sequence illustrated in Figure 1 or including the insertions or deletions of the amino acid sequence illustrated in Figure 3.

20 12. A NAALAD-ase III protein or a functional equivalent, derivative or bioprecursor thereof, encoded by the nucleotide sequence illustrated in Figure 4.

25 13. A protein according to claim 12 which is a human NAALAD-ase III protein.

14. A protein according to claim 12 or 13 having the amino acid sequence illustrated in Figure 4.

30 15. A NAALAD-ase IV protein or a functional equivalent, derivative or bioprecursor thereof, having an amino acid sequence encoded by the nucleotide sequence illustrated in Figure 5.

35 16. A protein according to claim 15 which is a human NAALAD-ase IV protein.

17. A protein according to claim 15 or 16 having the amino acid sequence illustrated in Figure 5.

18. A DNA expression vector which comprises a
5 cDNA molecule according to any of claims 1 to 8.

19. A vector according to claim 18 which vector comprises a cytomegalovirus promoter.

10 20. A vector according to claim 18 or 19 which comprises a sequence encoding a reporter molecule.

21. A host cell transformed or transfected with the vector according to any of claims 18 to 20.
15

22. A host cell according to claim 21 which cell comprises a prokaryotic or eukaryotic cell.

23. A host cell according to claim 21 or 22
20 which comprises a COS cell.

24. A transgenic cell, tissue or organism comprising a transgene capable of expressing a protein according to any of claims 10 to 17.
25

25. A transgenic cell, tissue or organism according to claim 24 wherein said transgene comprises a vector according to any of claims 18 to 20.

26. A cDNA or nucleic acid molecule according to any of claims 1 to 9, or a functional equivalent thereof, for use as a medicament.
30

27. Use of a cDNA or nucleic acid molecule
35 according to any of claims 1 to 9, or a functional

fragment thereof, in the preparation of a medicament in the treatment of neural diseases such as Alzheimer's disease, schizophrenia, ALS, Parkinsons disease, peripheral neuropathy, Huntingdon's disease,
5 acute brain injury, multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma, ischaemia or dementia.

28. A pharmaceutical composition comprising a
10 nucleic acid or cDNA molecule according to any of claims 1 to 8 and a pharmaceutically acceptable carrier, diluent or excipient therefor.

29. A method of determining whether a compound
15 is an inhibitor or an enhancer of activity of a NAALAD-ase protein according to any of claims 10 to 17 which method comprises contacting said compound with NAALAD-ase protein in the presence of [³H]N-acetyl-L-aspartyl-L-glutamate (NAAG), and monitoring for the
20 extent of hydrolysis NAAG compared to a control of said NAALAD-ase and NAAG which is not contacted with said compound.

30. A compound identified as an inhibitor or
25 enhancer of NAALAD-ase activity according to claim 29.

31. A compound according to claim 30 for use as a medicament.

30 32. Use of a compound according to claim 30 in the preparation of a medicament for the treatment of neural diseases or disorders such as Alzheimer's disease, schizophrenia, ALS, Parkinsons disease, peripheral neuropathy, Huntingdon's disease, acute
35 brain injury, multiple sclerosis, exposure to

neurotoxins, peripheral nerve trauma, ischaemia or dementia.

5 33. A pharmaceutical composition comprising a compound according to claim 30 together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

10 34. A method of identifying a compound which is an inhibitor or an enhancer of expression of a NAALAD-ase protein according to any of claims 10 to 17 which method comprises contacting a host cell, tissue or organism expressing said protein with said compound and monitoring the expression of said protein compared
15 to a control which comprises said cell expressing said protein but which has not been contacted with said compound.

20 35. A method according to claim 34 wherein said NAALAD-ase expressing cell comprises a host cell according to any of claims 21 to 23 or a transgenic cell, tissue or organism according to claims 24 or 25.

25 36. A method according to claim 35 wherein said monitoring step comprises monitoring for the expression of said reporter molecule.

30 37. A compound identified according to the methods of claims 34 to 36.

38. A compound according to claim 37 for use as a medicament.

35 39. Use of a compound according to claim 37 in the preparation of a medicament for treating neural

disorders such as Alzheimer's disease, schizophrenia,
ALS, Parkinsons disease, peripheral neuropathy,
Huntingdon's disease, acute brain injury, multiple
sclerosis, exposure to neurotoxins, peripheral nerve
5 trauma, ischaemia or dementia.

40. A pharmaceutical composition comprising a
compound according to claim 37 together with a
10 pharmaceutically acceptable carrier, diluent or
excipient therefor.

41. A method of quantifying enzyme activity
which method comprises
15 (a) contacting said enzyme with a radiolabelled
hydrolysable substrate thereof in the
presence of SPA beads;
(b) stopping said reaction by adding glycine
buffer; and
20 (c) monitoring the signal from said substrate
bound to said beads.

42. A method according to claim 41 wherein said
substrate comprises [³H]NAAG.
25

43. A method according to claim 41 or 42 wherein
said enzyme comprises a NAALAD-ase according to any of
claims 10 to 17.

GCCAAGAGTCCGAGGATGAGTGGACGAAGGTGTTGGGGCTGGGGCTGGGGCTGCTGCCCTCTTGGGGCTGGGGATCATCTCGGCCA 90
 M Q W T K V L G L G L G A A A L L G L G I I L G H
 CTTTGCCATCCCCAAAAAGCCAATCACTCGCCCCCAGGACCTGGACCTGGAGATCTCGAGACCGTCATGGGGCAGCTGGATGCCCA 150
 F A I P K K A N S L A P Q D L D L E I L E T V X G Q L D A H
 CAGGATCCGGGAGAACCTCAGAGAACTCTCCAGGGAGCCACACCTGGCCCTCCAGCCCTCGGGATGAGGACCTGGTGACGTGCTGCTGCA 270
 R I R E N L R E L S R E P H L A S S P R D E D L V Q L L L Q
 GCGCTGGAAGGACCCAGAGTCAGGCTGGACTCGGGCCGAGGCTNACGTAAGTGTGCTGCTCTCCCTAGCCAGGAGCAGCCCA 360
 R W K D P E S G L D S A E A X T Y E V L L S F P S Q E Q P N
 CGTCGTGGACATCGTGGGCCCCACTGGGGGCATCATCACTCTGCCACCGGACTGAGGAGAACGTGACCGGGAGCAAGGGGGCCAGA 450
 V V D I V G P T G G I I H S C H R T E E N V T G E Q G G P D
 TGTGGTACAACCCATGCTGCTATGCTCTCTTCTGGAACCCACAGGGGCTCTCTGCTATGCCAACCGGGCGCGGAAGAAGACTTTAA 540
 V V Q P Y A A Y A P S G T P Q G L L V Y A N R G A E E D F K
 GGAGCTACAGACTCAGGGCATCAAACTTGAAGGACCAATTCGCCCTGACTCGATATGGGGGTAGGGGCTGGGGCCAAAGGCTGTGAACGC 630
 E L Q T G I K L E T I A L T R Y G G V G R G A K A V N A
 TGCCAAAGCAGGGGTAGCTGGGGTGTGGTGTACACAGACCTGCGGACATCAAGATGGGCTGAGCTACCCGACGAAACCTTTCCAA 720
 A K H G V A G V L V Y T D P A D I N D G L S S P D E T F P H
 CTCTGTGTACCTGCCCCCTCAGGAGTGGAGCGAGGCTCTACTACGAGTATTTTGGGGACCTCTGACTGCTTACCTTCCAGCCGTCCC 810
 S W Y L P P S G V E R G S Y Y E Y F G D P L T P Y L P A V P
 CTCTCTCTCCGGTGGACCTTGCCAAATGCTCCGGATTTCGCCCAATTCCTACACAGCCCATGGGCTTCAGGATGCAAGAGACCTGCT 900
 S S F R V D L A N V S G F P P I P T Q P I G F Q D A R D L L
 CTGTAACTCAACGGAACTTTGGCCCCAGCCACCTGGCAGGGAGCACTGGGCTGCCACTACAGGTGGGTCCGGCTTCGGGCTGACGG 990
 C N L N G T L A P A T W Q G A L D G C H Y R L G P G F R P D G
 AGACTTCCAGCAGACAGCCAGGTGAATGTAGCGCTTACAACCGCTGGAGCTGAGGAACCTCTTCCAACGCTCTGGGATCATCCGTGG 1080
 D F P A D S Q V N V S V Y N R L E L R N S S N V L G I I R G
 GGCTGTGGAGCTGATGCTACGTGCTGTATGGGAACCAACGAGACAGCTGGGTGACCGGGCTGTGGACCCACAGTGGCAGCCCGT 1170
 A V E P D R Y V L Y G N H R D S W V H G A V D P S S G T A V
 CCTCTGGAGCTCTCCCGTGTCTGGGACCCCTGCTGAAGAAGGGCAGCTGGGCTCTCGCAGATCAATCGTGTTCGGAGCTGGGGGGC 1260
 L L E L S R V L G T L L K K G T W R P R R S I V F A S W G A
 TGAGGAGTTTGGGCTCATTTGGCTCCACGGAATTCACAGAAGAGTTCTTCAACAAGCTGAGGAGCCGACGGTGGCTACATCAACGTGA 1350
 E E F G L I G S T E F T E E F F N K L Q E R T V A Y I N V D
 CATCTCGGTGTTTGCACAGCTACCCCTTAGGGTGCAGGGGACGCCCTGTCTCAGAGCGTCTCTCTCTGCAACCAAGAGATCCGCTC 1440
 I S V F A N A T L R V Q G T P P V Q S V V F S A T K E I R S
 ACCAGGCCCTGGGACCTGAGCATCTACGACAACGGATCCGGTACTTCAACCGCAGCAGCCCGGTGACGGCTGGTCCCGAGCTGGG 1530
 P G P G D L S I Y D N W I R Y F N R S S P V Y G L V P S L G
 TTCTCTGGGTGCTGGCAGGACTATGCACCCCTTCGTTCACTTCTCGGCTATCTCTCCATGGACATTCCTATACCTATGACCGGAGCAA 1620
 S L G A G S D Y A P F V H F L G I S S M D I A Y T Y D R S K
 GACTTCAGCCAGGATCTACCCACCTACCAACAGCCCTTGAACCTTTGACTATGTGGACAAGTTTGGACCCGGGCTTCAGCAGCCA 1710
 T S A R I Y P T Y H T A F D T F D Y V D K F L C P G F S S H
 TCAGGCTGTGGCCCGCAGCAGCGGGAGTGTGATTCTCGGCTCAGTGACAGCTTCTTCTGCCCCCTCAAGTCAGTGACTACAGTGAGAC 1800
 Q A V A R T A G S V I L R L S D S F F L P L K V S D Y S E T
 ACTCCGAGCTTCTGAGGAGCCAGCAAGATCTTGGGGCCCTGCTGGAGCAGCAGCATCAGCCTGGGGCTCTGGTGACTGCACT 1890
 L R S F L Q A A Q Q D L G A L L E Q K S I S L G P L V T A V
 GGAGAAGTTTGAAGCAGAACTGACGCTTGGGCAACGATATCAACACTGAGAGGGGAGCCCTGACCCCTGACAGTCCGGATGCT 1980
 E K F E A E A A A L G Q R I S T L Q K G S P D P L Q V R M L
 CAATGACCAAGTTGATGCTCTTGGACGGACCTTTCTGAACCCTAGAGCCCTTCCAGAGGAACGCTACTACAGCCATGTGCTCTGGGACC 2070
 N D Q L M L L E R T F L N P R A F P E E R Y Y S H V L W A P
 TTGCGACGGCTCCGTAJTCACATTCCGGGGTATCCAAATGCTGCTCCAGGGCCAGGGACACAGCTTCTGGATCTGAAGCTTGGGCTGA 2160
 S H G L S H I P G L S N A C S R A R D T A S G S E A W A E
 GGTCCAGAGACAGCTCAGCATTTGTGTGACAGCCCTGGAGGGTGGGAGCCACCTGAGGCTGTGGCTGACCTCTGACCCAGCCCTC 2250
 V Q R Q L S I V V T A L E G A A A T L R P V A D L
 TTCTTCAGCCCTCCCTTACTCCGCTGCTTTATATTTACAAAGTCTTTGTGTTTTTAAAAJCTCTTT 2340

FIG. 2

human **LCCTKVLGLGLGAAALLGLGILGHFAIPKINSNP-----QDLLEILETVIGQLDASRIRENLRSLREPHLASSPRDEQLVQLLIQ** : 85
 rat **MFHAKILGVGIGAAALLGLGILGHFAIPKATZPLASSVSDSQDLLEILETVIGQLDASRIRENLRSLREPHVATSARDEQLVQLLIQ** : 90

human **RUKTPESGLDSAEAXTVEYLLSFPSOEOPHVDIVGPTGGIIRKCHRTEENVTEGCGGPDVVQPYAAYAFSGTFQGLVYANRGASEDTF** : 175
 rat **RUKDSASGLDTAKTYEDTVLLSFPSIEOPHSVEIVGPNSTVPIISFQPFKRLTGECAEPNVLOPYAAYAFPGTFKGLVYANRGASEDTF** : 180

human **ELCTAGIKLGGTIALTRYGVGRGAKAVNIAKRGVAEVLVYTDPADINDGLSSPDETFPNSVYLPSPGVERGSYYEYFGDPLTPYLPAMF** : 255
 rat **KLEAEGINIKGTIALTRYGVGRGAKAINAARHGVAEVLVYTDPADINDGLSPDETFPNSVYLPSPGVERGSYYEYFGDPLTPYLPAMF** : 270

human **SSFRVLLANVSGFPPIPTQPIGFQDARDLLCHLNGTLAFATUGGALGCHYRLGPGFRPDGDFPDSDQVNVSVYHRLRLNSSNVLGIIQG** : 335
 rat **VSRFLDPHVISGFPPIPTQPIGFEDAKILLCHLNGTSAPDSUQCALGCEYRLGPGFBPNGNFPACSEVNVSVYHRLRLNSSNVLGIIQG** : 350

human **AVEPDRYVLYGHERDSUVEGAVDPSSGTAVLLELSRVLGTLKKGITURPRRSIVFASUGAEFGLIGSTEFTEEFNMLQERTVATINVL** : 445
 rat **AVEPDRTVIYGHHRDSUVEGAVDPSSGTAVLLELSRVLGTLKKGITURPRRSIIFASUGAEFGLIGSTEFTEEFNMLQERTVATINVL** : 450

human **ISVFANATLFPQGTTPVQSVVFSATKEIRSPGPGDLSIYDNUIRYNRSSPVYGLVPSLGLGAGSDYAFVHFLGISSHDIAITYDRSF** : 535
 rat **ISVFANATLFPQGTTPVQSVVFSATKEIRSPGSSGLSIYDNUIRYNRSSPVYGLVPSHGLGAGSDYAFVHFLGISSHDIAITYDRSF** : 540

human **TSARIYPTYHTAFDITFDYVDFKLDPGFSSHQAVARTAGSVILRLSDSFPLPKVSDYSETLSFLQAAQDGLGALLECESISLGPLVTAV** : 625
 rat **TSARIYPTYHTAFDITFDYVDFKLDPGFSSHQAVARTAGSVILRLSDSFPLPLVSDYSETLSFLQAAQENLGALLESENISLGPLVTAV** : 630

human **EKFNDAALACRISTLOFSPDPLOVRMLNDQMLLEPTFLNPRAFPEERYYSHVLUAPSHGLRSHIPGLSHICSRRRTASGSEAVAF** : 715
 rat **EKFNDAALACRISTLOFSPDPLOVRMLNDQMLLEPTFLNPRAFPEERYYSHVLUAPNTASVATFGLNLYVAAQIINSGAEAVAF** : 720

human **VQRQLSIVDTALEGAAATLRPVADL--** : 740
 rat **VERQLSIAMVALEGAAATLRPVTDL--** : 745

FIG. 3

MQWTKVLGLGLGAAALLGLGIIIGHFAIPKKANSLAPQDLLEILETVMGQLDAHRIRENLRSLRPHLASSPRDEDLVOLLORWKC? 90

deletion 1 (bp 497-619)

ESGLDSAEAXTYEVLLSFPSQEQPNVWDIVGPTGGIIHSCRTEENVTEGQGGPDVVQPYAAYAPSGTPOGLLVYANRGAEDFKELOTO 130

GIFLEGTIALTRYGGVGRGAKAVNAAXEGVAGVLVYTPADINDGLSSPDETFPNSWYLLPPSGVERGSYYEYFGDPLTPYLPAVPSSF3V 270

gtgagcgtctacaaccgcctggagctgaggaaacctctccaaagctctcgsscat--a---
tggggctgtggaagcctggtagaccctctctctgtctgctcgcaacccagggccctt---
gctctggaagcgcgtctctctcatccagctccgctctgcccacacccagccca---
ctgcgccacgtctctctctctctctctggtctctctgcccctctctctctcgccca;

deletion 2 (bp 903-1007)

insertion at bp 1094

DLANVSGFPPIPTQPIGFQDARDLLCNLNGTLAPATWQGALGCHYRLGPGPRPDGDPADSQVNVSVYNRLRLNSSNVLGIIRGAVEP- 359

-----DRYVLYGNHRDSVWHGAVCPSSGTAVLLELSRVLGTLKK 199
 EFSSCCLHPRPLLCGCRCPHPALPLPPSPAPPAHLSLSSGSLPLPLWF

GTWRPRRSIVFASWGAEZFGGLIGSTEFTEFFNKLQERTVAYINVDISVFANATLRVOGTPPVQSVVFSATKEIRS?GPGDLSIYDNWIR 439

deletion 3 (bp 1525-1615); deletion 4 (bp 1525-1615)

gtgagggagggagacaaagggacatccctgagacacaggacaggagagagggctgaaagacctgagccctggcccttgccacaccttgccctgcag-

YFNRSSPVYGLVPSLGS LGAGSDYA²⁵FEVHFLGISSMDIAYTYDRSKTSA³¹IYPTYHTAFDTE³⁷DYVDKELDPGFSSHOAVARTAGSVILRL 579

RLQQPSGCGPDSGECDSPAQ*

RARLQPGS PPTTQPLTPLTMWTSFWTRASAAIRLWPGQRGV*

7 insertion at bp 1637

EEGDKGHPETRTCEAD*

gtatgcacagccctgacctgaggtatggggagccctgcaccccatgactgagccactgcttgctcctcacag

Insertion at bp 1870

SDSFFFLPKVSDYSETLRSFLOAACDCLGALLEOHSISLGPLVTAVEKFEAEAAALGORISTLQKGSPPDLOVRMLNDOLMLLERTFINP 619

GMHSPDPEVWGALHPHD*

RAFP EERYYS HVLWAPSEGLRSEI PGLENACSRARDTASGSEAWAEVOECLSIVVTALEGA AATLRPVADL* 740

FIG. 4

CTCAAGAAGCCATGGCGGAATCCAGGGCCGCTGTACCTTTGGATGTGCTGGCTGCGCTGGCATCTTTCTGTATGGATTTATGG 30
M A E S R G R L Y L W H C L A A A L A S F L M G F M V
TGGGCTGGTTTATTAAGCCTCTCAAAGAAACAACCACTTCTGTGCGCTATCATCAAAGTATACGGTGGAACTGGTATCCGAAATGAAAG 160
G W F I K P L K E T T T S V R Y H Q S I R W K L V S E H K A
CTGAAAACATCAATCTTCTGCTTTTACAAAGCTTCTCATCTGCGAGGAACAGAACAAATTTCTTGTGCGCAAGAAAATCC 270
E N I K S F L R S F T K L P H L A G T E Q N F L L A K K I C
AAACCCAGTGGAAAGAAATTTGGACTAGATTACGCAAGTTGGTTCAATTATGATGCTCTTATCTTACCCCAATGAGACAAATGCCAACT 360
T Q W K K F G L D S A K L V H Y D V L L S Y P N E T N A N Y
ATATATCGATTGTGATGAACATGAACTGAGATTTTCAAACATCATACCTTGAACCAACCAGATGGCTATGAGAAATGTTACAAATA 450
I S I V D E H E T E I F K T S Y L E P P P D G Y E N V T N I
TTGTGCCACCATATAATGCTTTCTCAGCCCAAGGCATGCCAGAGGAGATCTTGTATATGTGAACATGCTGCGACTGAAGACTTTTTC 540
V P P Y N A F S A Q G M P E G D L V Y V N Y A R T E D F F K
AACTAGAAAGAGAGATGGGCATCACTGTACTGGGAAGATTGTTATTGCAAGATATGGAATAATCTTCAGAGGAAATAAGTAAATAATG 630
L E R E M G I N C T G K I V I A R Y G K I F R G N K V X N A
CCATGTTAGCAGGAGCCATAGGAACTATCTGTACTCAGATCCAGCTGACTACTTTGCTCTGAGGTACAGCCATATCCCAAAGGATGGA 720
H L A G A I G I I L Y S D P A D Y F A P E V Q P Y P K G W N
ATCTTCTGGAAGTGCAGCCAGAGAGAAATGTGTTAAATTTGAATGGTGGTGGTACCCTCACTCCAGGCTATCCAGCAAAAGAAAT 810
L P G T A A Q R G N V L N L N G A G D P L T P G Y P A K E Y
ACACTTTCAGACTTGATGTTGAAGAAGGAGTGGGAATCCCCGAAATACCTGTATCATCCCATGGATATAATGATGAGAAATATTATTAC 900
T F R L D V E E G V G I P R I P V H P I G Y N D A E I L L R
GCTACTTGGGAGGAATGCTCCACAGATAAGAGTTGGAAGGGAGCCCTTAATGTGAGTTATAGTATCGGACCTGGCTTTACAGGAGAGT 990
Y L G G I A P P D K S W K G A L N V S Y S I G P G F T G S C
ATTCTTTCAGGAAGGTAGAAATGCACTGTTTATAACATCAATAAAATACAAAGGATTTACAATGATGTTGGAATATCAGAGGATCTGTG 1080
S F R K V R M H V Y N I N K I T R I Y N V V G T I R G S V E
AACCTGACAGGTATGTTATCTGGGAGGTACCCGGGACTCCTGGGTATTTGGAGCTATTGACCCCAACCACTGGGGTTGCTGTTTTCGCAAG 1170
P T G T A I L G G H R D S W V F G A I D P T S G V A V L Q E
AAATGCGCGGAGTTTGGAAAACGTATGAGTAAAGGCTGGAGACCTAGAAGAACTATCATTTTTCGAGCTGGGATGAGAGAAATTTG 1260
I A R S F G K L M S K G W R P R R T I I F A S W D A E E F G
GACTTCTGGGTTCCACAGAAATGGGCTGAGGAGAAATGCAAAATACCTCCAGGAGAGAAGCATTTGCTTATCAACTCGGATTCATCTATAG 1350
L L G S T E W A E E N V K I L Q E R S I A Y I N S D S S I E
AAGGCAATTATCTCTCAGAGTTGACTGTACTCCCTCTTTACCAATTAGTGTATAAATGACAAAAGAGATCCCGAGCTTGTATGATG 1440
G N Y T L R V D C T P L L Y Q L V Y K L T K E I P S P D D G
GGTTTGAGAGTAAATCACTGTATGAAAGCTGGTTGGAAAAAGACCTTCACCTGAAATAAAAAATTTGCTTAGAATCAATAAGCTGGGAT 1530
F E S K S L Y E S W L E K D P S P E N K N L P R I N X L G S
CTGGAAGTGACTTTGAAGCTTATTTTCAGAGACTTGGAAATGCTTCAGGAGAGCCCGTTACACTAAGAATAAGAAAACAGATAAGTACA 1620
G S D F E A Y F Q R L G I A S G R A R Y T K N K K T D K Y S
GCAGCTACCCAGTGTACACACAATTTATGAGACATTGAATTTGGTAGAGAAATTTATGACCCACATTTAAAAACAACCTTTCTGTGG 1710
S Y P V Y H T I Y E T F E L V E K F Y D P T F K K Q L S V A
CTCAATTACGAGGAGCACTGGTATATGAGCTTGTGGATTCTAAAAATCATCTTTTAAATATTCAAGACTATGAGAGGCTTTGAAAAAT 1800
Q L R G A L V Y E L V D S K I I P F N I Q D Y A E A L K N Y
ATGCAGCAAGTATCTATATCTATCTAAGAAACATGATCAACAATTAACAGACCATGGAGTATCATTTGACTCCTTATTTCTGCTGTGA 1890
A A S I Y N L S K K H D Q Q L T D H G V S F D S L F S A V E
AAAACTTCTCAGAGGCTGCTCAGATTTTCAATAACGACTTATACAAGTTGATCTTAACAATCCCATTCAGTGAAGATGATGAATGAC 1980
N F S E A A S D F H K R L I Q V D L N N P I A V R M M N D C
AACTGATGCTCCTGGAAGAGCAATCATGATCCTCTTGGTTTACCAAGGAAGCTGTTCTATAGGCACATCATTTTGTGCAAGTAGCC 2070
L M L L E R A F I D P L G L P G K L F Y R H I I F A P S S H
ACAACAAATATGCTGGAGAAATCATTTCTGGAATCTATGATGCTATCTTTGATATTGAAAATAAGCCAACCTCGTTTGGCTGGAAAG 2160
N K Y A G E S F P G I Y D A I F D I E N K A N S R L A W K E
AAGTAAAGAAACATATTTCTATGCGAGCTTTTACAATTAAGCAGCAGCAGGAACCTGAAAGAAATATTAGAAAGGCTCAAGTGGCT 2250
V K K H I S I A A F T I Q A A A G T L K E V L

FIG. 4 (cont.)

AGCCATTAAAGGTGTTGCTAAAAGTCGAGGATAAAATTCACCTTTCTGATAACTTATGAAGCCAGGGTGTCTAAACTCTTTTCATGTC 234:
ATGTTTTGATTATAGGCTTTGGTCCTTCATCTGCAAAAGCCTTTTCTTTTGTCTCTTAAAAGTTAATAATTATATTAGCAAGTGT 243:
AATCTAATGAAGTAAAAACTCTGTGTGGCAGAAAGTAAAGAAAAATCCCTAAATTATAGCAAGGAACATGAATTCTCAGACATTGTG 252:
AGTGTGGGAATGTAAAATGGTAAATCACTTTTGAAAAACAGTTTGGCAGTTTCTATAAAGTTAAACATACACTTTTACTTTAGGACTCC 261:
AGAATTCACCTTCTAGTTATTTATTCAGAGAAGGAAAAACAATGATCAGCAATACTTGTATGCATGTTCAATGCAACTTAAAAGCGT 270:
AAAAACCCCAATGTCCATCCACAGACGAATGTATAAACTGTGGTATCCATTACACAATAGACTACTTACTACTCAGCAATAAAAATGAA 279:
GTAACCTTCAATAAATGCAATATTATTCGGCAGACATTGTTGAAGGAAAAAGCCAGACAAACAACATACATAAAATATGTTTCTATTAGA 288:
TGAAGTGGCAAACTAATCTGTAGTGTAAAAATTAGATTAGTGATTGCCCTGGGCCAAGTGGCAGGTTGGGGAGGATGGCTGCAAGGAAGT 297:
ATGAGGAACTTTCTCCAATAGATGAGAAATTTCCGTATCTTGATCTGAGTGGCAAAATGTAAACTTAAATATATATAAAATTTTATGA 306:
AAGAAAATTAAGCCTCAATAAACGTGATTATAAAAAAAAAAAAAAAAAAGG 315:

FIG. 5

CGCGCGGAGGGCCCGAGCCAGTCAGGGGTGTGGCCGCCGCCACCGTAAGGCTAGGCCGAGCTTAGTCTGAGCGCCCTCCGTCG 90
CCGCCGTGAGAGCCGCCCTATCAGATTATCTTAACAAGAAAAACCACTGGAAAAAAATGAAATTCCTTATCTTCGCAATTTTCGGTGG 180
M K F L I F A F F G G
TGTTCACTTTTATCCTGTGCTCTGGGAAAGCTATATGCAAGAAATGCCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAGAAATAGC 270
V H L L S L C S G K A I C K N G I S K R T F E E I K E E I A
CAGCTGTGGAGATGTTGCTAAAGCAATCACTAACCTAGCTGTTTATGGTAAAGCCAGAACAGATCTCTATGAGCGATTGGCACTTCGGT 360
S C G D V A K A I I N L A V Y G K A Q N R S Y E R L A L L V
TGATACGTTGGAGCCAGACTGAGTGGCTCCAAGAACCTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGGA 450
D T V G P R L S G S K N L E K A I Q I M Y Q N L Q Q D G L E
GAAAGTTCACCTGGAGCCAGTGAGAAATACCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAAGAAATCAAGATAGC 540
K V H L E P V R I P H W E R G E E S A V M L E P R I H K I A
CATCTCGGGTCTGGCAGCAGCATTTGGACTCTCCAGAAAGCATTACAGCAGAAAGTTCTGGTGGTGACCTCTTCGATGAACGACAGAG 630
I L G L G S S I G T P P E G I T A E V L V V T S F D E L Q R
AAGGGCCTCAGAAGCAAGAGGGAAGATGTTGTTTATAACCAACCTTACATCAACTACTCAAGGACGGTGCAATACCGAACGCGAGGGGC 720
R A S E A R G K I V V Y N Q P Y I N Y S R T V Q Y R T Q G A
GGTGGAAAGTGCCAAAGGTGGGGCTTTGGCATCTCTCATTGATCCGTTGGCTCTCTTCCATCTACAGTCTCTACACAGGTATTACAGGA 810
V E A A K V G A L A S L I R S V A S F S I Y S P H T G I Q E
ATACCAGGATGGCGTCCCAAGATTCCAACAGCCTGTATTACCGTGGAAAGATGCAGAAATGATGTCAGAAATGGCTTCTCATGGATCAA 900
Y Q D G V P K I P T A C I T V E D A E M H S R M A S H G I K
AATTGTCTTACAGCTAAAGATGGGGGCAAGACCTACCCAGATAGTATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCC 990
I V I Q L K H G A K T Y P D T D S F N T V A E I T G S X Y P
AGAACAGGTTGTACTGGTCACTGGACATCTGGACAGCTGGGATGTTGGGAGGGTGCCATGGATGATGGCGGTGGAGCCTTTATATCATG 1080
E Q V V L V S G H L D S W D V G Q G A H D D G G G A F I S W
GGAAGCACTCTCACTTATTAAAGATCTTGGGCTGCGTCCAAGAGGACTCTGCGGCTGGTCTCTGGACTGCAGAAAGCAAGGTGGAGT 1170
E A L S L I K D L G L R P K R T L R L V L W T A E E Q G G V
TGGTGCTTCCAGTATTATCAGTTACACAAGGTAATATTTCCAACTACAGTCTGGTATGGAGTCTGACGAGGAACCTTCTTACCCAC 1260
G A F Q Y Y Q L H K V N I S N Y S L V M E S D A G T F L P T
TGGGCTGCAATTCACTGGCAGTGAAGGCAAGGCGCATCATGGAGGAGGTTATGAGCCTGCTGCAGCCCTCAATATCACTCAGGTCTCT 1350
G L Q F T G S E K A R A I M E E V M S L L Q P L N I T Q V L
GAGCCATGGAGAGGGACAGACATCAACTTTGGATCCAAGCTGGAGTGCCTGGAGCCAGTCTACTTGATGACTTATACAAGTATTCTT 1440
S H G E G T D I N F W I Q A G V P G A S L L D D L Y K Y F F
CTTCCATCACTCCACGAGACACCATGACTGTATGGATCCAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGT 1530
F H H S H G D T M T V M D P K Q H N V A A A V W A V V S Y V
TGTTCAGACATGGAAGAAATGCTGCTAGGTCTTAGAAAACAGTAAGAAAGAAACGTTTTCATGCTTCTGGCCAGGAACTCTGGGCTGTC 1620
V A D M E E M L P R S *
AACTTTGGAAAACCTCTCTTCAATAACAAATTCATCCAATTCACTCTCAAAGCACAACCTTATTTCATGCTTCTGTTTATTATCTTTCT 1710
TGATACTTTCCAAATTCCTGATTCTAGAAAAAGGAATCATTCTCCCTCCCTCCACCATAGAAATCAACATATGGTAGGGATTACAG 1800
TGGGGCATTTCTTTATATCACTCTTAAAAACATTGTTCCACTTTAAAGTAAACACTTAATAAATTTTGGAGATCTCTG 1884

FIG. 6

NAALAD I	MWNLLHETDSAVATARRPRWICAGAVLA--GGPFLGFLFGWFIKSSNEATNTPKHNK-APDSTKRENTFRFYNTQIPHHRTGEQ	68
NAALAD II	-----MQNTKVLGLGGAALGLGIIILGHFAIPKKANSLAP-----QDLDLELETVYGIDNHRITDEMARELSREPHLSPP	75
NAALAD III	-----MAESRRLYLWMCDAALASFLGFMVGVFIKPLKETTTSRYHQSHRWKLVSSEKRENTFRFYNTQIPHHRTGEQ	78
NAALAD IV	-----MKFLIFAPPGEVHLISLCSGRKICKNCKSKRTPEEIKKEIASCGDIAKATINLAVYGKAQNRSYDRLALLVDIVGERUSSEKN	83
NAALAD I	FFQAKQIQSQNR--PFGANGVEIAHYVNTSYENKTHPEVYSINEDGNETPNTSLFPPPEPYNVSDIAPESASPOSMPEGDIVY	176
NAALAD II	DEDFVQLRLORNDPESQISAEAXTGEWANSPPSDEQPHVVDIYVGTGGIIMHSCHRTENVTSCGGPDIYQPAALAPSGTESMAY	185
NAALAD III	FFLAKRKIQQLR--KPSLSAKVHVDVLEETENETNAKYSISYDEHETEIPKTYLEPPEDEGVNVTITVPEANISACSMPEEDIVY	186
NAALAD IV	LEKAIQIHYQNLQ--EDGTEKVIPEVRPHWGERGESAVMIEPRINKIAILGLGSIGTPEES-----	185
NAALAD I	VNYARTDFEKTERRDKINCSCRTIAPYGVFGRGNVVKHLOLACAKVILLYSDADYCAPGYSK--MDESNMGGGVORGMTNLNG	263
NAALAD II	ARRG-AEEDRELOTQCKLE-THATPACGGRGCAVHAKKSVAGLVVNDERINDGLSPDETENAVMEPSSGVVERGSYYEY--	262
NAALAD III	VNYARTDFEKTERRDGINCCRTIAPYGVFGRGNVVKHMLAELCHILYSDADYCAPGYSK--MDESNMGGTAAORGMTNLNG	263
NAALAD IV	-----TFAELIVTSPDEORRASE-ARGK-----IUVVNCQYINISRTVOY--RTQSAVEAAKVGALASIASVA	268
NAALAD I	AGRETTGEGYENMEYARRGAAVAPESIEVHEIGTYDACKNEK*GESAPDSSTRGSKVPTANVGESEPTGNF---STQKRYMHMSTN	350
NAALAD II	FGDEHTFENYENVPSSRVDTANVSPEPIETQEHIECOARDLCCNINETLAE-AT*QSAUGCHALGPPRRZDGDPPADSSQNNSSVYNR	341
NAALAD III	AGRETTGEGYENMEYARRGAAVAPESIEVHEIGTYDACKNEK*GESAPDSSTRGSKVPTANVGESEPTGNF---STQKRYMHMSTN	340
NAALAD IV	SFSIYSEHTGIGQDQD-----EIERETACTITVEDADMSRMASHGIKIVIQIRMGAKTYPTDS-----	368
similar to bacterial Zn ²⁺ dependent peptidase		
NAALAD I	EWTRINMHTETGCAWEDPMTILGCHDSM--PCHQPCSAKWHMIVPSPTTKKER-PRRRRTILPAS-DDEPPILGSTPABG	437
NAALAD II	EHANSSHULEIIEFAVEPDRWLYGNHDSM--VAVDPSSTAVILLEISRWLCTLLKKEI-PRRRRTIPAS-DDEPPILGSTPABG	429
NAALAD III	KUTRIYHIVETLGSVEPDRWLYGNHDSM--VAVDPSSTAVILLEISRWLCTLLKKEI-PRRRRTIPAS-DDEPPILGSTPABG	427
NAALAD IV	-----ENTAEETESKYSEQWVLYSGHDSMDVCGGAKHDGGA--FISWEALSILKDL-LRPRRTILVLTADEPCVGAFTY--	348
NAALAD I	NSHLLQERGVAYTFASSIEGYLQANDCHSLYSYMHNLHEPKSEDEPEGKSLYESWTKKSPSEFSGIRAMSKLGSNDSEVVFQR	527
NAALAD II	PPNKTGPPNYATLWHSIPFAHARVOCSPVQSUPFSAHETIRGEGPDLSDYDMWIRYFNRSQSVYGLVESHGSGTASSTAPVHF	519
NAALAD III	NVMTAERSYATLWHSIEGYLQANDCHSLYSYMHNLHEPKSEDEPEGKSLYESWLEKDPSEENKMERHNNKSGSDZALVQR	517
NAALAD IV	LHVNINISNYSLVKEAGTFLPTGQPTGS-----	378
NAALAD I	FGTASGRARVWETNKPSCVSLHFWYDYSLVENFYDEMKEYHTVAQVRGQVSELANSVIHPDCRUIAVVIRKADKIYSIMK	617
NAALAD II	FGTASMDIAPVDRSKTSARLYETTHAFEDYVDKFLHGGSSHOAARTACVLLRLSDSFFMLKLSQDSETHSSLOAACODLGA	609
NAALAD III	FGTASGRARVWETNKPSCVSLHFWYDYSLVENFYDEMKEYHTVAQVRGQVSELANSVIHPDCRUIAVVIRKADKIYSIMK	607
NAALAD IV	-----EKARATMCEMMLLCLNITQLSHGEGTDNEMHQAQVPGASLDELKYKYFFHMSHGDTAVM	443
NAALAD I	HPQIKTYSVSDSDSFFPKKKEITLISKPSERTQDFDKS--NDVIANNDQDMFMBAL*SLGLDRPARRVATYAGSHNKYAGEZF	705
NAALAD II	---LSEQHSILGPIVITDEKFEAEAAALGQPISTLQKGSPODQVRMMDQPLHETPEHRAFPSEERTSHVYIAPSHG---LASHI	693
NAALAD III	HDCCTDHGVSVDSPFAKHSAAASDPFKK*IQVDLN--DPSANTHMDQMLALRA*SLGLDRPARRVATYAGSHNKYAGEZF	695
NAALAD IV	DPKQNVAAAVAVSYVVAQNDMLPRS-----	472
NAALAD I	EGYVLLFOIESKVDPSKIGGKQVYAAFTCAAEWISKA--	750
NAALAD II	EGYVLLCSFARDTASGSEVYVQVQSHVVTAHGAARPRMADL	740
NAALAD III	EGYVLLIFDIENKANSRLAFYKQVYAAFTCAAEWISKA--	740

FIG. 7

Growtree Phylogram of: noaladase_short_distances, Tree Tree_1
June 17, 1998 17:46

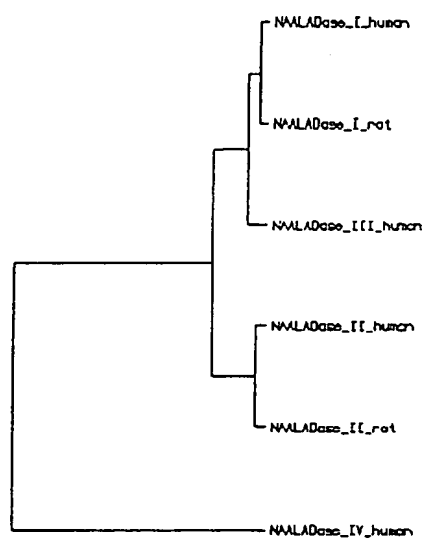


FIG. 8

```

NAALAD I   DAQKLLEKMGGSAPPD-SSWRGSLKVPNNVGPFTGNF---STQKVKLHIHSTNEVTRIYNVIGTLRGAVEPTRYVILGG 376
NAALAD II  DAEILRYLGGIAPPD-KSUKGALNVSYSIGPGFTGSD---SFRKVRHVYNINKITRIYIVVGTLRGAVEPTRYVILGG 386
NAALAD II  --RDLCNLNGTLAP--ATUQGALGCHYRLGPGFRPDGDFPADSQVNVSVYNRELNRSSIVLGIIRGAVEPTRYVILGN 397
NAALAD IV  SPHTGHOEYQDGVPKIPTACITVEDAEHMSRMASHGIK---IVIQLKGAKTYPDTDS-FYIVAEITCSKYPEQVVLGG 389
APE 3 yeast -----TKHTVATVGVPYKVGKKLIANIALNIDYSLYFAMDSYVEFIKTONIADTKHG-DPTNIVLALGA 313
P96152     QITNTIRALSSFNRFYTTASGAQASDLANEVRLIS---SLPGSRLEQIKHSGYNQ-KSVVLTICSEKPEEVLVGG 399
AMPX vibpr QITGTISLSEFTNRFYTTTSGAQASDIASEUQALSA---SLPNASYKQVSHSGYNQ-KSVVHTITSEAPATDRIKGG 312
APX Strgr  -----NNGGNR--AHGRPGYKASVDYVKAKLDA---GYTTTQQFTSGGATG-YNLIANUPGG-DPTKVLALGA 34

```

```

*          *          *
NAALAD I   HRDSU-VF-----G-TI-DPQSG-LAV-VHEIMVSFGTL-KKEG-WRPRTILFASDDAEFFGLLGSTETA-EE-NS 439
NAALAD II  HRDSU-VF-----G-AI-DPTSG-LAV-LQELINSFGKL-HSKG-WRPRTILFASDDAEFFGLLGSTETA-EE-NV 429
NAALAD II  HRDSU-VH-----GAV-DPSSGTDVL-L-ELSEVLGTLK-KGTURPERSIVFASGGAEEFFGLLGSTET-EEFTN 432
NAALAD IV  HLDSDVV-----GCGAMDDGGG-LFISU-EALSLI---KDLG-LRPKRTLRLVLGTAEPCGGGAFCHY-QLHEV 382
APE 3 yeast HSDS-VEE-----GPGINDGSGTISL-L-NVAFQATH-----FRINNKVRFADGAEEFFGLLGSTETAYNLTK 375
P96152     HLDST-LGSHINEQSIAPGADDDASGIRSL-S-EIDFVL---RDNN-FRPKRSALMAFAEEVGLRGSODPA-NQYKA 270
AMPX vibpr HLDST-IGSHINEQSVAPGADDDASGIRAV-T-EIDFVL---SENN-FQPKRSIAFMATAEEVGLRGSODLA-NQYKS 273
APX Strgr  HLDV-VSS-----GAGINDNGGSLAV-L-ETALAV---SRAG-YQPDHARFAMGAEEFFGLLGSTETKFFV-NNLPS 147

```

```

*
NAALAD I   RLLOERGVAYNNAUSSI-EGNYTLRVDCIPLHYSLVENLKEK-KSPDEGFEGKS-YESUTKKS--PSPFSGHPRISK-L 516
NAALAD II  KILQERSIAYINSSSI-EGNYTLRVDCIPLLYQLVYKLKEK-PSPDGFEKSK-YESGLEKD--PSPENKNLPRINK-L 516
NAALAD II  KL-QEPTVAYINVDISV-FANATLRVQCTPPQSVVFSALKERSPGPGD--LSIYDNGIRYFNRSSPVYGLVPSLGS-L 518
NAALAD IV  NIS--NYSLVRESAGT-FLPTGLQFTGSEKARA---IHEEMH---SLQPLNITQ-----VLSEK 414
APE 3 yeast ENSKIR--VFEDYDMA-SPNYEYETIDANNKENP--KGSSEK-----NLYVDYKAK-----HLNYTLVPFDG 435
P96152     QGK--KVVSVLQLDNTNYRGSIEDIVFIDYDTS---NLQFPT-----TIDENLPKL-----TYG---YDRCG 317
AMPX vibpr EGK--NVVSALQLDNTNYKGSAQDVVFIIDYDTS---NFQVPT-----QNDENLPKL-----TYG---FDTCC 311
APX Strgr  AD-RSELAGYLNFDH-IGSPNPGYFYVDDDPVIEK--TFKNYFAG-----LNVPTETETE-----GDGRSDHAPFFN 211

```

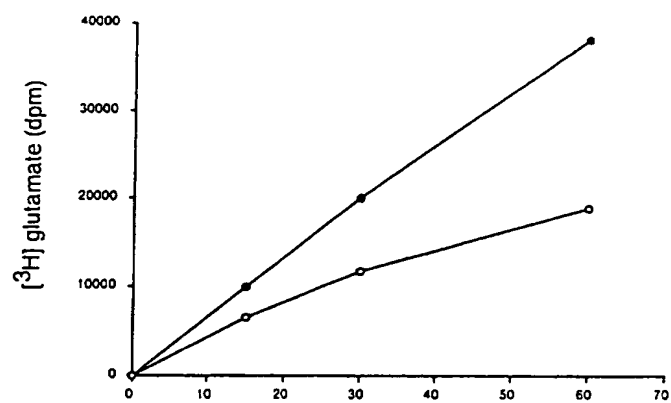
```

*
NAALAD I   SGNDFEVFFQRLGSIASGRARYTNNWETNKFSGYPLYESVYETVELVEK---FYDPMKYH-LTVLQVRGG----- 553
NAALAD II  SGSDFEAYFFQRLGSIASGRARYTNKKTDKYSSYPVYETIYEFELVEK---FYDPTKKQ-LSVQLRGA----- 572
NAALAD II  AGSDYAPFVHFLGISSHDIATYDRSKTSARIYPTIETAFDFDYDK---FLDPCSSH-QAVRTAGS----- 574
NAALAD IV  EGTDIIN-FEIQACVPGASLLDLYKYFF-----FRHSHGDTHTVSDPKQNVAAAVAVVSVVVDMEHMLPRS----- 473
APE 3 yeast RSDYVGFINNGIPAGGIATGAENNVNNGKVLDRCYEQLCDVSNLSWDAFITNTKLIASHSVATYIDSEFGFPKRETQKH 515
P96152     YACSDHASLHKAFSAAMPFESKFDYN-----PKIEMSDTLANSPT--GNHAVTETKLGLAYVIENAN----- 391
AMPX vibpr YACSDHASLHNAGYPAAMPFESKFDYN-----PRIETODTLANSPT--GSHAKMTQLGLAYVIENG----- 394
APX Strgr  VGVVUGGLTGAICYTKSAAQAQGGGTAGQAFDRCYHSCDSLSNINDTALDRNSDAAAHAIWTLSSGTGEPPT----- 394

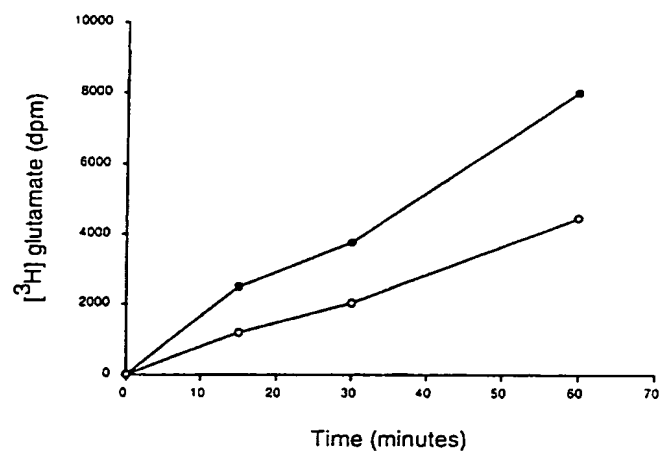
```


FIG. 9

A



B



C

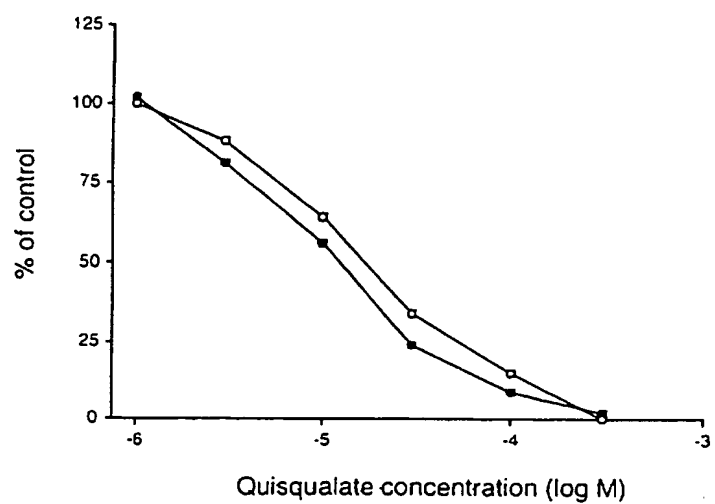
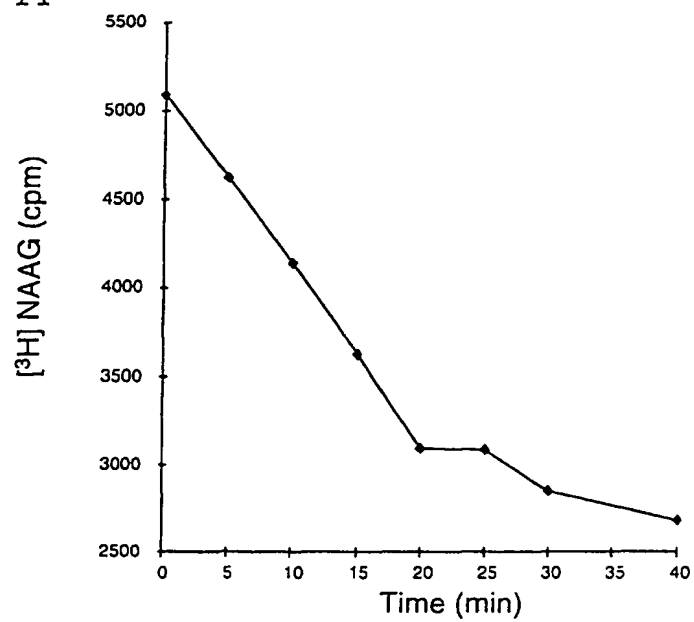
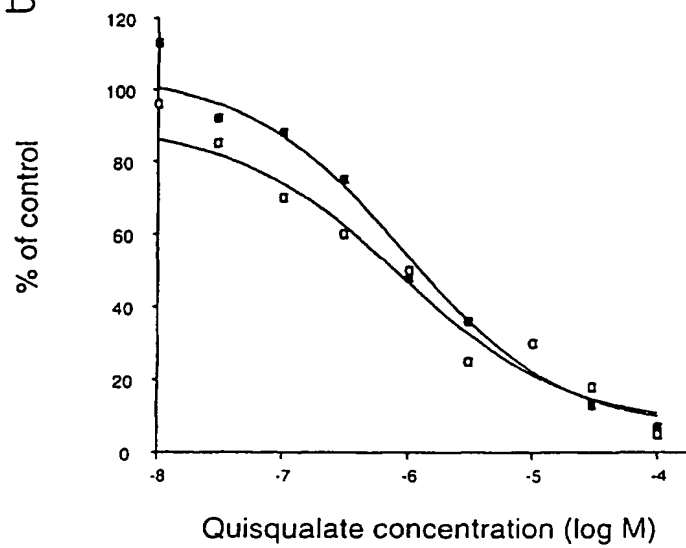


FIG. 10

A



B



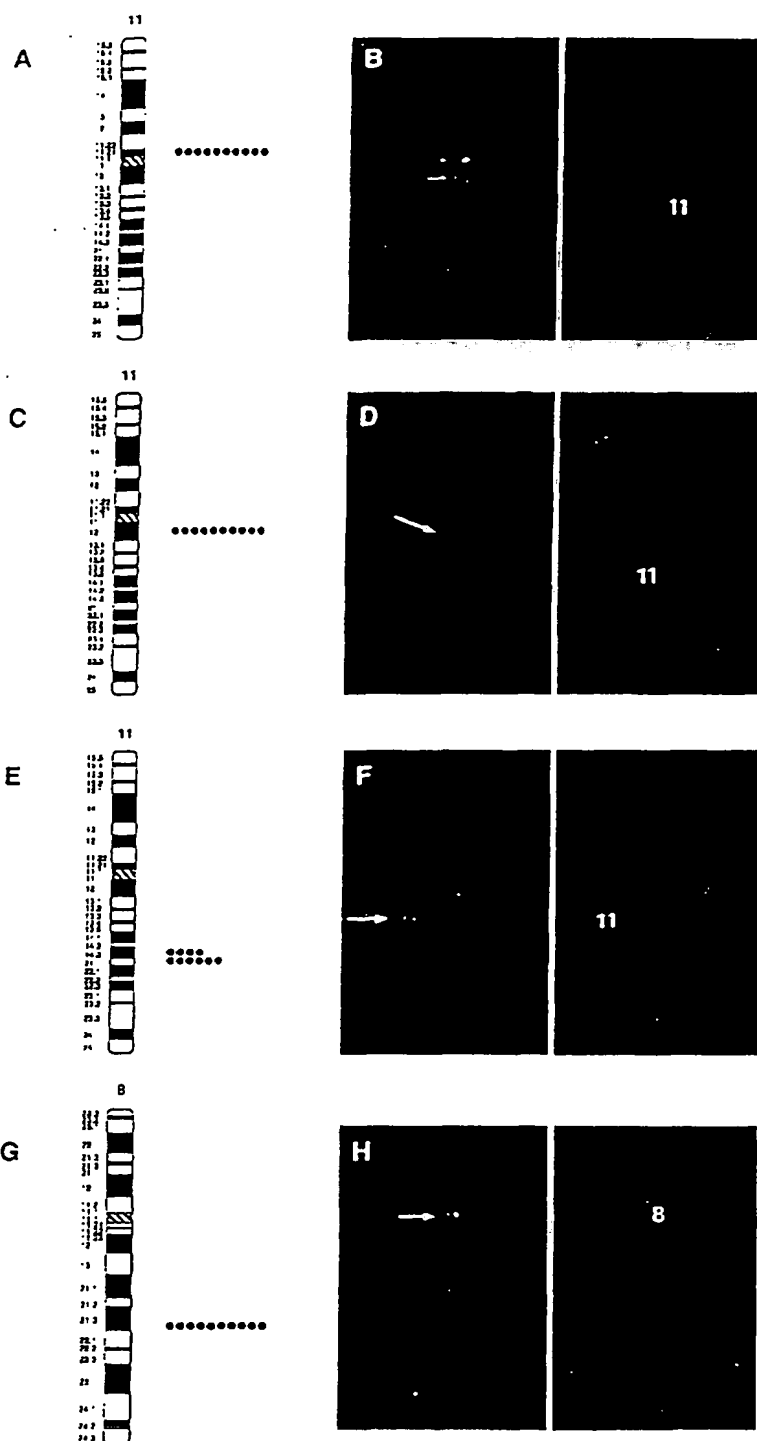
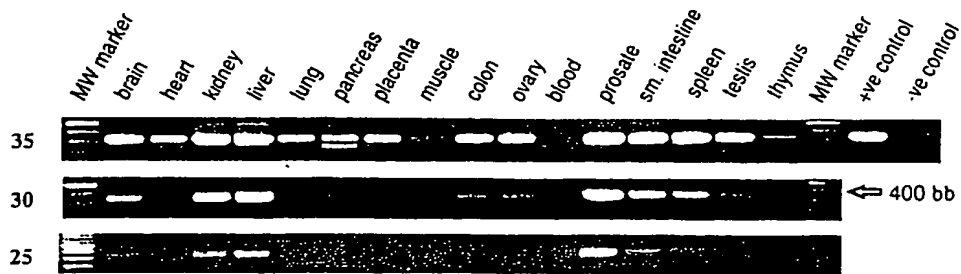
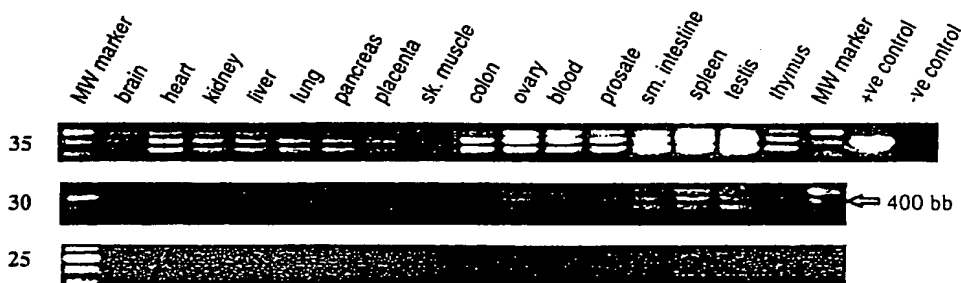


FIG. 13

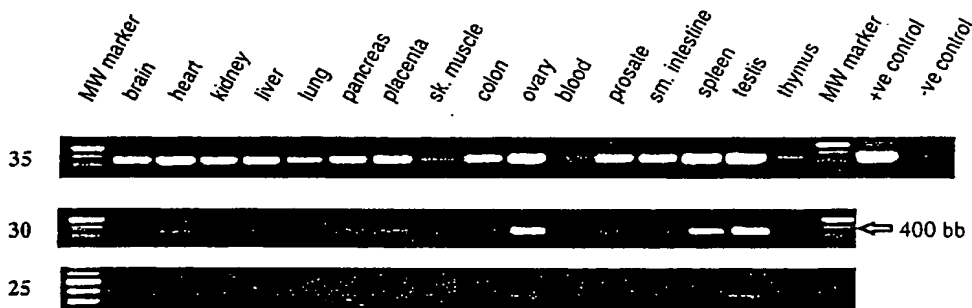
A



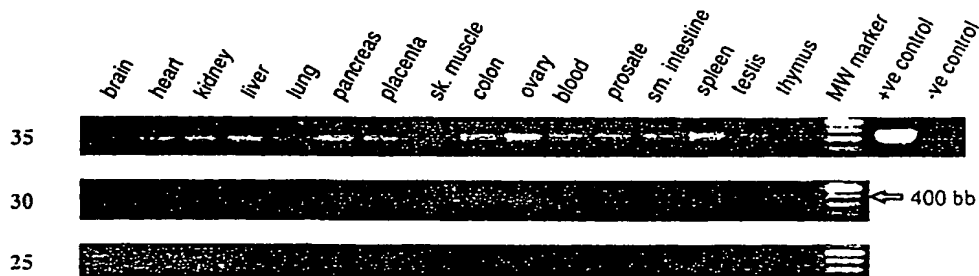
B



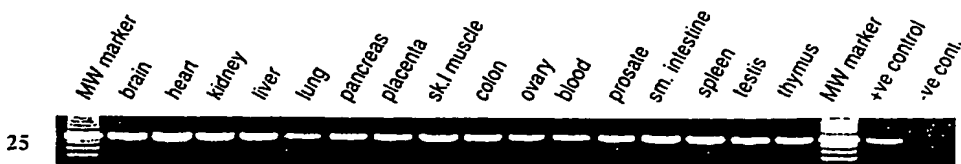
C



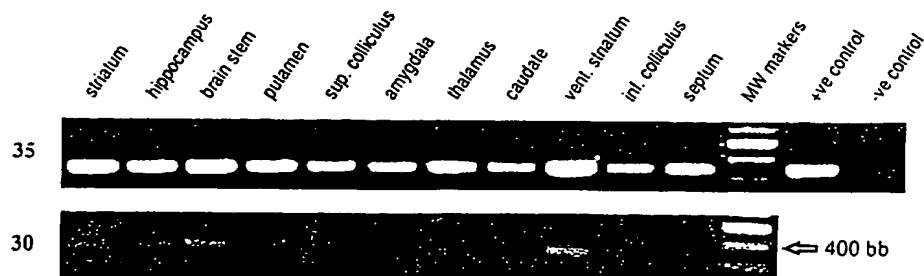
D



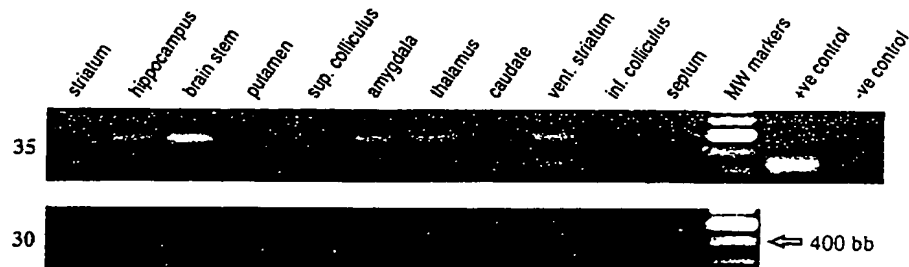
E



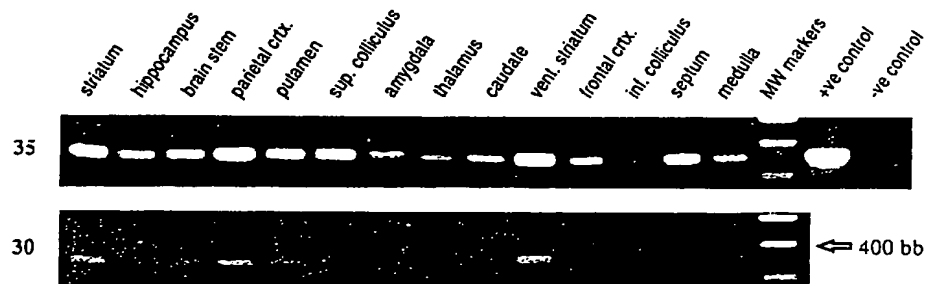
A



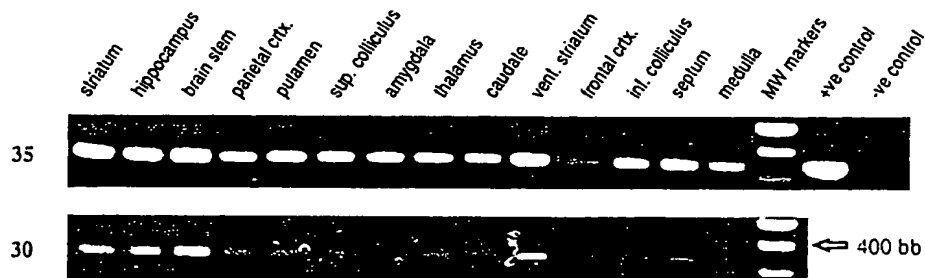
B



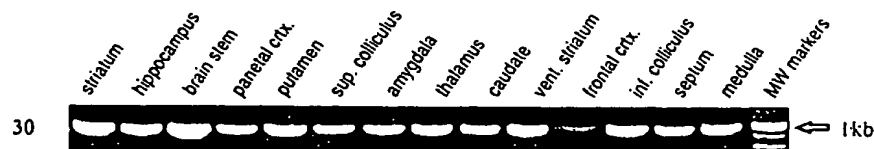
C



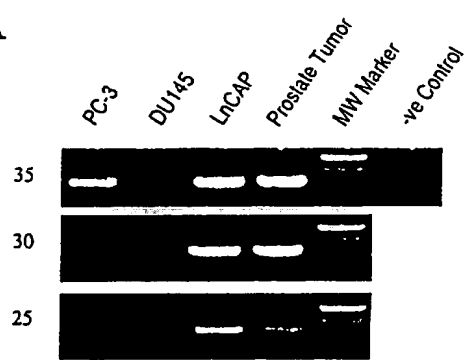
D



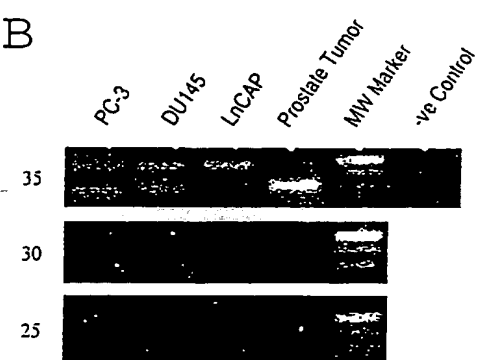
E



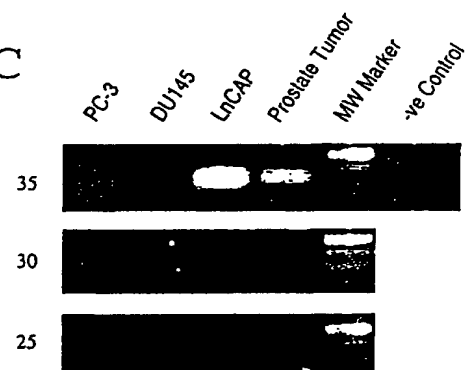
A



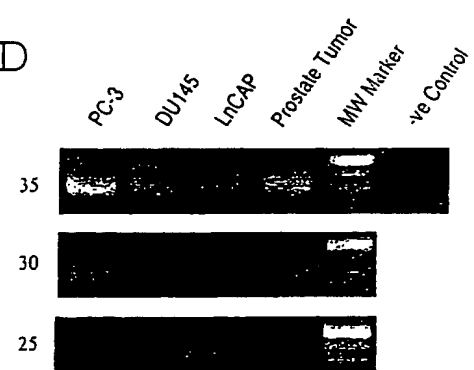
B



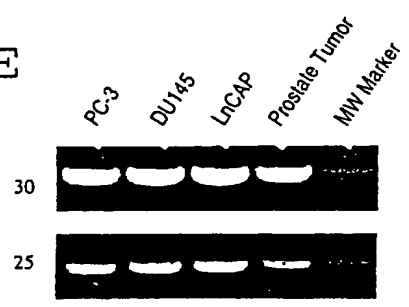
C



D



E



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.